

Open Research Online

The Open University's repository of research publications and other research outputs

The identification and analysis of mutation in the Cockayne Syndrome B gene

Thesis

How to cite:

Mallery, Donna Louise (1999). The identification and analysis of mutation in the Cockayne Syndrome B gene. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 1999 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:
<http://dx.doi.org/doi:10.21954/ou.ro.0000e27e>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

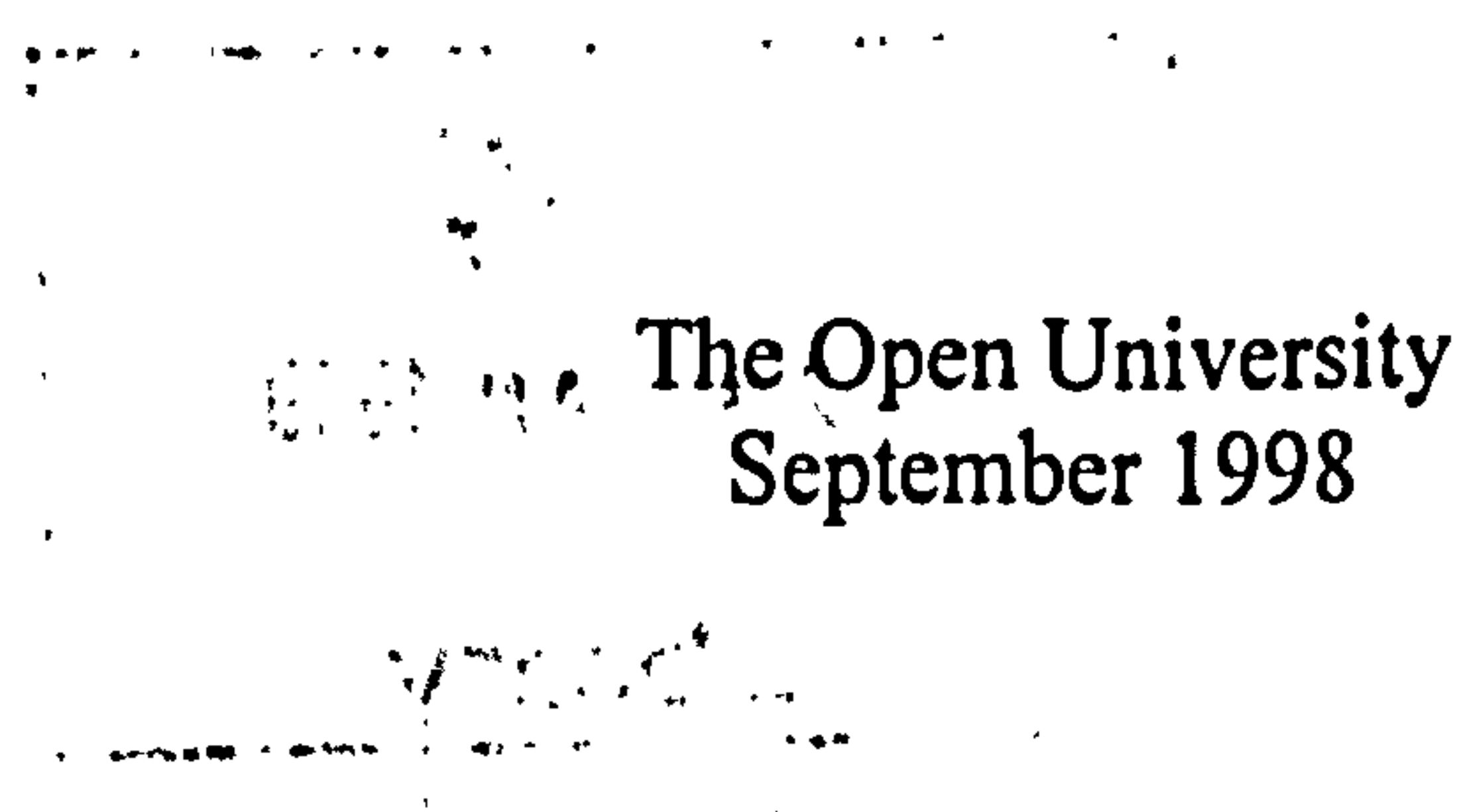
UNRESTRICTED

The identification and analysis of mutations in the Cockayne syndrome B gene.

A thesis submitted for the degree of Doctor of Philosophy in Life Sciences.

By

Donna Louise Mallery BSc



Author number: P927 737X
Date of award: 22 January 1999

DECLARATION

The work presented in this thesis is my own except where otherwise stated and has never been submitted to any other university for any other degree.

D L Mallery

**Donna Louise Mallery
September 1998**

ACKNOWLEDGEMENTS

I would like to thank everyone, past and present at the CMU for their patience, support and friendship over the past four years. Thanks also go to our colleagues in Rotterdam and Pavia for helpful discussions.

I would especially like to thank Alan Lehmann for his help, supervision and support, particularly when times were difficult. Also to Tracy, for not knowing when to keep quiet.

A big thank you to Rick, it's been hard but it is finally finished.

ABSTRACT

Cockayne Syndrome (CS) is a rare autosomal recessive disorder characterised by neurodegeneration, dwarfism and at least three of the following; hearing loss, dental caries, pigmentary retinopathy, characteristic facial appearance and photosensitivity. Cells from CS patients fail to recover RNA synthesis after irradiation and exhibit a loss of transcription-coupled repair, with overall genome repair being unaffected. There are two genetic complementation groups of CS alone, A and B, with the majority of patients belonging to group B. The genes defective in each of the complementation groups have been cloned, the *CSA* gene in 1995 and *CSB* in 1990.

For the purposes of this study the *CSB* gene was sequenced in patients from complementation group B, in an attempt to identify the causative mutations. The analysis of thirteen patients from different backgrounds has revealed a wide variety of mutations in the *CSB* gene. A considerable number of the mutations found in CS-B patients resulted in severely truncated products. Several patients possessed two alleles affected in this way and it is unlikely that any functional protein is produced, confirming that *CSB* is a non-essential gene. The mutations identified did not reveal any regions within the gene that could be termed as hotspots. There was, however a tendency for the mutations to be located towards the 3' two thirds of the gene, indicated by the clustering of the mutations in this region. The severity of the mutation does not however correlate with the site or type of mutation.

Clustering of the mutations towards the 3' end and the high levels of conservation in the central part of the gene prompted a study into the functional significance of the N- and C-terminal ends of the protein. Also, the presence of a highly acidic region of amino acids and a stretch of glycine residues led to a study of the effects of removing and replacing these regions. Removal of the glycine domain results in non-functional protein with respect to cell survival after UV irradiation, whereas the removal of seven glutamic acid residues from the acidic rich region, does not appear to have a particularly dramatic effect. Deletion of the C-terminal 25 amino acids of *CSB* totally destroys the repair ability of the gene. In contrast, cDNAs deleted at the N-terminus are able to at least, partially retain repair activity.

CONTENTS

ACKNOWLEDGEMENTS

ABSTRACT

LIST OF TABLES AND FIGURES

ABBREVIATIONS

AIMS

CHAPTER 1 GENERAL INTRODUCTION	1
1.1 Nucleotide excision repair	1
1.1.1 Nucleotide excision repair in <i>E.coli</i>	1
1.1.2 Nucleotide excision repair in eukaryotes	2
1.1.3 Nucleotide excision repair disorders	3
1.1.3.1 Xeroderma pigmentosum	3
1.1.3.2 Cockayne syndrome	4
1.1.3.3 Trichothiodystrophy	4
1.1.4 Genes involved in NER	4
1.1.4.1 XPA/RPA	4
1.1.4.2 XPC/HHR23A&B	8
1.1.4.3 XPG	9
1.1.4.4 ERCC1-XPF	10
1.1.4.5 TFIIH	12
1.1.4.6 TTD-A	14
1.1.4.7 XPE	14
1.1.5 Repair synthesis	14
1.2 Basal transcription by RNA polymerase II	15
1.2.1 TFIIH	17
1.3 Transcription-coupled repair	20
1.3.1 Transcription-coupled repair in <i>E.coli</i>	21
1.3.2 Transcription-coupled repair in <i>S.cerevisiae</i>	24
1.3.2.1 Yeast TCR mutants	25
1.3.2.2 Double mutants	25
1.3.3 Mammalian transcription-coupled repair/Cockayne syndrome	28
1.3.2.3 CSA	31
1.3.2.4 CSB	31
1.3.2.5 SWI/SNF complexes	32
1.3.2.6 CSB knockout mice	34
1.3.2.7 CSB and transcription	34
1.3.2.8 RNA polymerase II ubiquitination	37
1.3.2.9 Transcription-coupled repair-a link with mismatch repair	38
CHAPTER 2 MATERIALS AND METHODS	40
2.1 Cell culture techniques	40
2.1.1 Primary human fibroblasts	40
2.1.2 CHO cells	40
2.1.3 Sub-culturing and harvesting	40
2.1.4 Transfection	40
2.1.5 Freezing	41
2.1.6 UV survival-Method I	41

4.3.2	CS1ABR	66
4.3.3	CS1TAN and CS2TAN	66
4.3.4	25627	67
4.3.5	CS3TAN and CS1IAF	67
4.3.6	CS1BE, CS4BR and CS2BI	68
4.3.7	Polymorphisms	69
4.4	Functional complementation	70
4.5	Discussion	72
CHAPTER 5 CS2GO AND CS3GO		75
5.1	Introduction	75
5.2	Materials and methods	75
5.3	Results	76
5.4	Functional analysis	79
5.5	Discussion	80
CHAPTER 6 DELETION ANALYSIS		83
6.1	Introduction	83
6.2	Construction of deleted cDNAs	84
6.2.1	C-terminal deletions	84
6.2.2	N-terminal deletions	85
6.2.3	Internal deletions	85
6.3	Results	86
6.3.1	C-terminal deletions	87
6.3.2	N-terminal deletions	87
6.3.3	Glycine and acidic domain deletions	88
6.4	Discussion	88
CHAPTER 7 GENERAL DISCUSSION		90
7.1	Mutational analysis	90
7.2	Western blots	95
7.3	Future work	98
7.3.1	Functional analysis of the acidic and glycine domains	98
7.3.2	Expression and purification of recombinant proteins	99
7.3.3	SNF2-like mutations	99
7.4	Models for the role of CSB	100
7.4.1	A model for the role of CSB	104
7.5	Concluding remarks	105
REFERENCES		106
APPENDIX I Primers		
APPENDIX II Alignment of human, mouse and yeast CSB proteins		
APPENDIX III Full sequence of the <i>CSB</i> cDNA		

LIST OF TABLES AND FIGURES

Table 1.1	A summary of the human proteins and their <i>S.cerevisiae</i> counterparts which are involved in NER.
Table 2.1	The cDNAs, PCR primers and amplification conditions used for the PCR amplification of the <i>CSB</i> cDNA.
Table 3.1	The PCR cycling conditions employed for amplification of the <i>CSB</i> cDNA.
Table 4.1	The origin, clinical features and mutations of patients.
Table 4.2	The mutations and polymorphisms found in CS-B patients.
Table 7.1	The nature of the mutations created in CSB and the analagous mutations in SNF2 from which they were derived.
Figure 1.1	Diagrammatic representation of the <i>E.coli</i> nucleotide excision repair mechanism.
Figure 1.2	Picture representation of the mechanism of mammalian NER.
Figure 1.3	The FEN1 family of proteins.
Figure 1.4	A diagrammatic representation of the CSB protein indicating putative functional domains.
Figure 1.5	The SWI/SNF family of proteins.
Figure 3.1	The location of the PCR primers used to amplify the <i>CSB</i> cDNA.
Figure 3.2	Formamide agarose gel of a standard total RNA preparation.
Figure 3.3	Agarose gel of a control PCR product from cDNAs primed by oligo dT and random hexamers.
Figure 3.4	Agarose gel showing amplification of the 5' end of <i>CSB</i> .
Figure 4.1	RNA synthesis after UV irradiation in CS-B and wild-type cells.
Figure 4.2	Diagrammatic representation of CSB illustrating the position and nature of mutations found in CS-B patients.
Figure 4.3	Sequence demonstrating the absence of exon 10 in CS1MA.
Figure 4.4	Sequence showing the presence of both the wild-type and single T deleted alleles from position 3614 in CS2BE.
Figure 4.5	Sequence showing the presence of both the wild-type and the 25bp insertion from intron 13 in CS1ABR.

- Figure 4.6** Sequence showing the presence of both the wild-type sequence and the 26bp insertion at position 3686 in CS1ABR.
- Figure 4.7** Sequence showing the presence of both the T and C tracks at position 2282 in 25627 and the homozygous change present in CS1TAN.
- Figure 4.8** Sequence showing a homozygous G to A change at position 1630 in CS2TAN.
- Figure 4.9** Sequence showing the presence of both nucleotides C and T at position 1436 in 25627.
- Figure 4.10** Sequence of wild-type (panel a) and patient CS3TAN (panel b) around position 2630 showing a T in place of A in panel b.
- Figure 4.11** Sequence showing the nucleotide change from T to G at position 2630 in CS1IAF.
- Figure 4.12** Sequence showing the presence of both the C and T nucleotides at position 2087 in CS1BE.
- Figure 4.13** Sequence showing both the wild-type and single A deleted alleles at position 3615 in CS2BI.
- Figure 4.14** Sequence showing both C and T are present at position 29 in CS4BR.
- Figure 4.15** Sequence showing both C and T present at position 3204 in CS2BI.
- Figure 4.16** Graph illustrating the ability of *CSB* cDNAs with the mutations R670W, W851R and V957G to correct the UV sensitivity of UV61.
- Figure 4.17** Graph illustrating the ability of *CSB* cDNA carrying the mutation P1042L to correct the UV sensitivity of UV61.
- Figure 5.1** Sequence showing the presence of both T and C nucleotides at position 2630 in CS2GO.
- Figure 5.2** Sequence showing the presence of both alleles from patient CS2GO at position 213, one with each of a C or T nucleotide.
- Figure 5.3** Sequence showing the presence of both G and A nucleotides at position 3368 in CS2GO, and just an A in CS3GO.
- Figure 5.4** Sequence showing a change from G to A at position 4317 in CS2GO and CS3GO.
- Figure 5.5** Sequence showing the presence of both G and A nucleotides at position 3716 in CS2GO, and just an A in CS3GO.
- Figure 5.6** Sequence showing the presence of both the full length and deleted alleles in patient CS3GO.

- Figure 5.7** Graph illustrating the ability of *CSB* cDNA with the mutations W851R and L45P to correct the UV sensitivity of UV61.
- Figure 5.8** Graph illustrating the ability of *CSB* cDNA carrying the mutation Q1413R to correct the UV sensitivity of UV61.
- Figure 5.9** Graph illustrating the ability of *CSB* cDNA carrying the mutation R1213G to correct the UV sensitivity of UV61.
- Figure 5.10** Graph illustrating the ability of *CSB* cDNA carrying the mutation V1097M to correct the UV sensitivity of UV61.
- Figure 5.11** Diagrammatic representation of *CSB* illustrating the position and nature of mutations found in the CS-B patients CS2GO and CS3GO.
- Figure 6.1** Diagrammatic representation of *CSB* showing the putative functional domains and the deleted cDNAs.
- Figure 6.2** The cloning strategy for the creation of N-terminal deletions of *CSB* cDNA.
- Figure 6.3** The 39 amino acid sequence of the acidic domain of *CSB*.
- Figure 6.4** Survival curves for clones derived from transfection of UV sensitive UV61 cells with C-terminally truncated *CSB* cDNA.
- Figure 6.5** Survival curves for clones derived from transfection of UV sensitive UV61 cells with N-terminally truncated *CSB* cDNA.
- Figure 6.6** Survival curves for clones derived from transfection of UV sensitive UV61 cells with *CSB* cDNA deleted for the seven glutamic acid residues 378-384.
- Figure 6.7** Survival curves for clones derived from transfection of UV sensitive UV61 cells with *CSB* cDNA deleted for the glycine domain.
- Figure 7.1** All mutations found in the *CSB* gene during this study of patients with Cockayne syndrome type B.
- Figure 7.2** Western blots to test the sensitivity of a commercially bought monoclonal HA antibody, monoclonal HA purified from 12CA5 hybridomas and polyclonal *CSB* antibody.
- Figure 7.3** Western blot of cell extracts from cells transfected with wild-type or mutant *CSB* cDNA.
- Figure 7.4** Western blot with cell extracts from untransfected UV61 cells and cells transfected with a mutant *CSB* cDNA.
- Figure 7.5** Western blot of cell fractionation from cells wild-type and mutant for *CSB*.
- Figure 7.6** Diagrammatic representation of *CSB* showing the position and nature of the SNF2-like mutations created.

Figure 7.7 Survival curves for clones derived from transfection of AA8 cells with mutant *CSB* cDNAs.

ABBREVIATIONS

AAF	N-2-acetylaminofluorene
AF	N-2-aminofluorene
BPDE	benzo(a)pyrene diol epoxide
CAK	cyclin-dependent activating kinase
Cdk	cyclin-dependent kinase
CHO	chinese hamster ovary
CPD	cyclobutane pyrimidine dimer
CS	Cockayne Syndrome
DMSO	dimethyl sulphoxide
DTT	dithiothreitol
EDTA	diaminoethanetetra-acetic acid, disodium salt
ERCC-	excision repair cross complementing
FEN1	flap endonuclease 1
GGR	global genome repair
HA	influenza haemagglutinin protein
HEPES	N-(2-hydroxyethyl) piperazine-N'-(2-ethanesulphonic acid)
HNPCC	hereditary nonpolyposis colon carcinoma
IPTG	isopropyl- β -D-thiogalactopyranoside
MCS	multiple cloning site
MEM	minimal essential medium
MFD	mutation frequency decline
NA-AAF	N-acetoxy-2-acetylaminofluorene
NER	nucleotide excision repair
1-NOP	1-nitrosopyrene
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
RFC	replication factor C
RNAPII	RNA polymerase II
RNAPII-CTD	RNA polymerase II – C-terminal domain
RPA	replication protein A
RT	reverse transcription
SDS	sodium dodecyl sulphate
SDW	sterile distilled water
SRB	suppressor of RNA polymerase B
TAF	TBP associated factors
TBP	TATA-box binding protein
TCR	transcription-coupled repair
TE	tris-HCl pH 8(10mM), EDTA (1mM)
TFII-	RNA polymerase II transcription factor
TRCF	transcription repair coupling factor
TTD	trichothiodystrophy
UDS	unscheduled DNA synthesis
UV	ultraviolet
XP	xeroderma pigmentosum

AIMS

This project is intended to provide insight into the function of the CSB protein and expand knowledge of Cockayne syndrome and how its characteristic symptoms arise. Such information will also contribute to the field of DNA repair, particularly when linked to transcription, thereby increasing knowledge of how mammals respond to DNA damage.

CHAPTER 1 GENERAL INTRODUCTION

1.1 Nucleotide excision repair

In order to maintain the genetic integrity of the genome, a number of mechanisms have evolved to effect repair of damage in DNA occurring both spontaneously and environmentally. One of the most significant and widely studied sources of environmental DNA damage is UV light, which causes lesions such as cyclobutane pyrimidine dimers and pyrimidine-pyrimidone (6-4) photoproducts. The major pathway involved in the repair of UV damage is nucleotide excision repair (NER) (Friedberg et al.,1995).

Generally, NER consists of five steps :

- a) damage recognition,
- b) incision of the damaged strand on both sides of the lesion,
- c) excision of the lesion-containing oligonucleotide,
- d) synthesis of a new strand,
- e) ligation.

1.1.1 Nucleotide excision repair in *E. coli*

In total, six gene products are required for NER in *E.coli* and these proteins alone can be used to reconstitute the reaction *in vitro* on naked DNA (reviewed in Friedberg et al., 1995). The mechanism of the reaction is outlined in figure 1.1 (Hoeijmakers, 1993).

The first step in the reaction is dimerisation of the UvrA gene product. UvrA is a DNA-dependent ATPase, and the hydrolysis of ATP is required for the dimerisation reaction and also increases the specificity of dimer binding to damaged DNA. The dimer once formed recruits a single molecule of UvrB, also in an ATP dependent manner. This complex has

been demonstrated as having a weak 5' to 3' helicase activity. It is thought that this activity may allow the complex to translocate along the DNA and scan for lesions.

When a lesion is encountered in the DNA, UvrB becomes bound at this site and the UvrA dimer is dissociated. The binding of UvrB is thought to result in a conformational change in the DNA helix, which is recognised and allows binding by UvrC.

The position of the 3' incision is variable but generally occurs four or five phosphodiester bonds from the lesion. In contrast the 5' incision always occurs eight bonds from the damage site. ATP is required for the incisions but is not hydrolysed. This suggests that it binds and causes a conformational change which allows the incisions to occur, effected by UvrC 5' and UvrB 3' to the damage.

The next few stages of NER are catalysed by the concerted actions of UvrD (DNA helicase II) and DNA polymerase I. It is suggested that, firstly, UvrD releases UvrC and the lesion containing oligonucleotide. DNA polymerase I binds the exposed 3' OH group and then displaces UvrB during synthesis of the new strand. The nicks remaining after repair synthesis are sealed by DNA ligase.

1.1.2 Nucleotide excision repair in eukaryotes

NER in eukaryotic cells follows the same basic steps observed in *E.coli*, but a far greater number of gene products are involved in the reaction. Much of the information gained on the mechanism of NER in humans has been provided by the study of mutant cell lines. In recent years NER has been reconstituted *in vitro* with purified proteins both for the yeast and human systems (Aboussekhra et al., 1995; Guzder et al., 1995). This has provided great insight into the proteins necessary for the reaction and identified accessory roles for others.

The reaction is highly conserved between *S.cerevisiae* and mammals with many proteins showing not only functional homology, but also great similarity at the amino acid level.

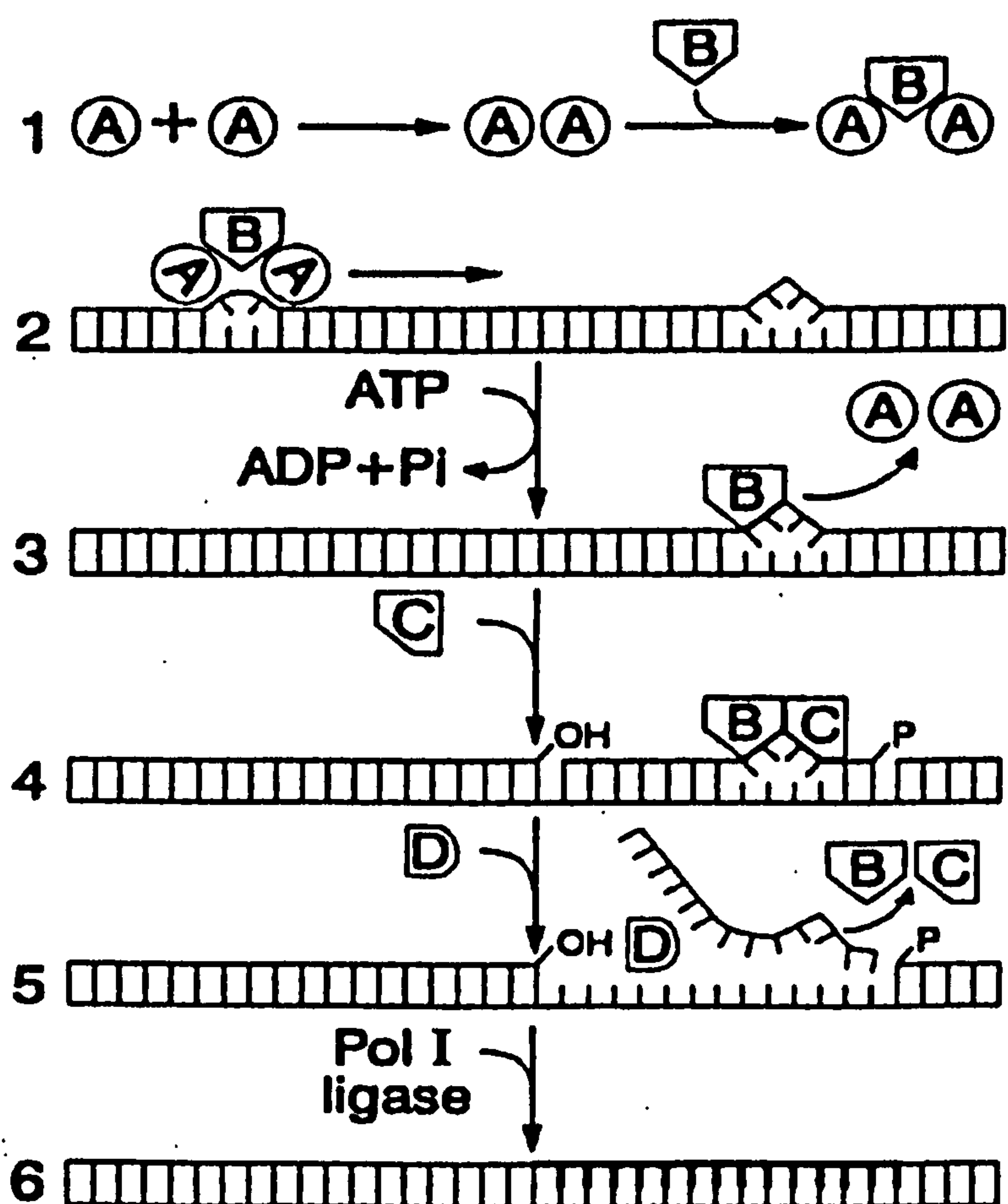


Figure 1.1 Diagrammatic representation of the *E.coli* nucleotide excision repair mechanism (Hoeijmakers, 1993).

Human Protein	Yeast Protein	Comments
XPA	Rad14	Binds damaged and single-stranded DNA
RPA	Rpa	Three subunits, single-stranded binding protein, binds XPA
XPC/hHR23A,B	Rad4/Rad23	Binds damaged and single-stranded DNA Not required for all excision repair substrates
TFIIH	TFIIH	Six to nine subunits including helicases XPB and XPD
XPG	Rad2	3' DNA endonuclease
XPF/ERCC1	Rad1/Rad10	5' DNA endonuclease
PCNA	PCNA	Stimulates DNA synthesis by polymerases δ and ϵ
DNA Polymerases δ and ϵ	Pol δ and Pol ϵ	Synthesises new DNA
DNA Ligase I	Lig I	Seals nicks

Table 1.1 A summary of the human proteins and their *S.cerevisiae* counterparts which are involved in NER, and the functions that they are likely to carry out in the reaction.

Table 1.1 outlines the essential factors required for NER in mammalian cells and their yeast counterparts.

1.1.3 Nucleotide excision repair disorders

Three major disorders are associated with defects in nucleotide excision repair, xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy.

1.1.3.1 Xeroderma pigmentosum

Xeroderma pigmentosum (XP) is a rare autosomal recessive disorder, with the most characteristic and easily identifiable symptoms being hyperpigmentation and sun-sensitivity (Kraemer and Slor, 1984). XP patients are in particular associated with a greatly elevated risk of skin cancers. In rare cases of XP these features can be accompanied by the clinical features of CS. Cells derived from XP patients demonstrate an increased sensitivity to killing after treatment with DNA damaging agents and in particular UV irradiation (Arlett and Lehmann, 1979). In addition to the increased killing, cells also demonstrate an increased mutation frequency in response to damage.

XP was first identified as a NER disorder in 1968 when Cleaver showed that patients with XP were unable to carry out NER after UV irradiation (Cleaver, 1968). Cellular diagnosis of XP patients is carried out by measuring the incorporation of ³H-thymidine into UV irradiated non-S phase cells by autoradiography, or into non-dividing cells by scintillation counting. These assays measure the levels of repair synthesis or unscheduled DNA synthesis (UDS). UDS experiments on XP cells demonstrated a large variability between different XP cell lines (Bootsma et al., 1970) and subsequently seven complementation groups, A to G were identified (Cleaver and Kraemer, 1989). The corresponding genes have all been cloned and their products participate in the different steps of NER. A further complementation group of XP has also been identified, termed XP variant, in which there is no defect in NER (Cleaver, 1972). XP variant cells have been shown to be defective in

post-replication repair (Lehmann et al., 1975) although no gene has as yet been identified in this group.

1.1.3.2 Cockayne Syndrome

Cockayne syndrome (CS) is a rare autosomal recessive disorder characterised by neurodegeneration and dwarfism as well as at least three of the following; hearing loss, dental caries, pigmentary retinopathy, characteristic facial appearance and photosensitivity (reviewed in Nance and Berry, 1992). Cells from CS patients exhibit a loss of transcription-coupled repair (see below), with overall genome repair being unaffected. The cellular and molecular characteristics of CS will be discussed in more detail later.

1.1.3.3 Trichothiodystrophy

Trichothiodystrophy is a rare autosomal recessive disorder characterised by sulphur-deficient brittle hair, ichthyosis, and mental and physical retardation (Pollitt et al., 1968). Patients also have distinctive facies with protruding ears and a receding chin. Some patients with TTD are photosensitive and cells derived from these patients are defective in NER (Stefanini et al., 1986), whereas those TTD patients without photosensitivity have no NER defect (Stefanini et al., 1987).

TTD has three complementation groups, two of which are the same as the XP complementation groups XP-B and XP-D, with the mutations occurring in the same genes, (discussed in greater detail below). The third group termed TTD-A consists of a single patient. The repair defect in cells derived from this patient can be corrected by fusion with cells from all NER complementation groups (Stefanini et al., 1993).

1.1.4 Genes involved in NER

1.1.4.1 XPA/RPA

XP-A patients exhibit one of the severest forms of XP suffering from elevated skin cancer levels and progressive mental retardation. Cells derived from patients in the XP-A

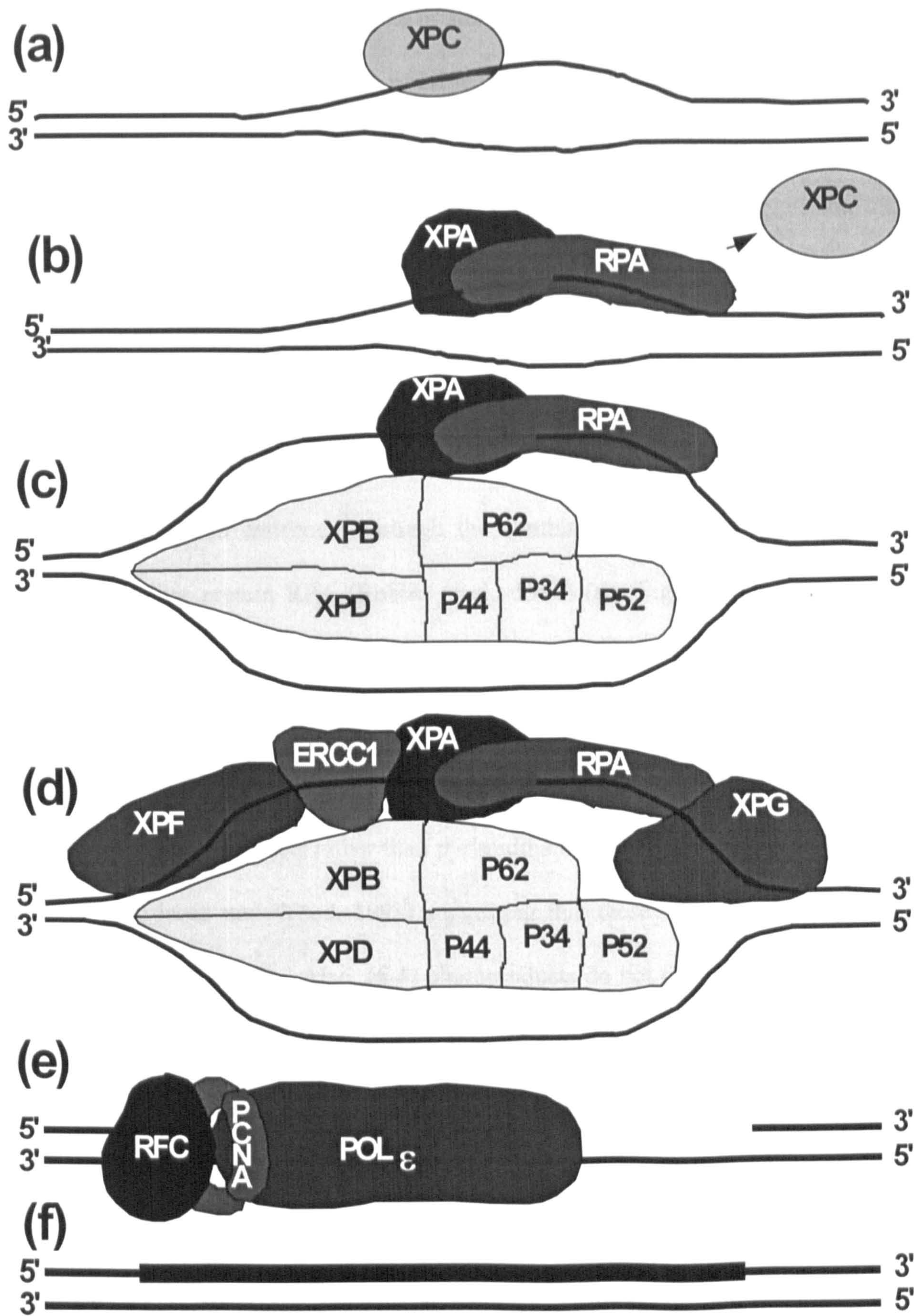


Figure 1.2 Picture representation of the mechanism of mammalian NER (adapted from Lehmann, 1995).

complementation group show no gene or strand specific repair (Evans et al., 1993) and are totally deficient in the incision step of NER (Reardon et al., 1993). Mutation studies have shown some Japanese patients to have reduced levels of mRNA and 16 from 21 patients studied possess the same mutation (Tanaka et al., 1990). In general, the mutations identified tend to correlate well with the clinical features observed (Satokata et al., 1990; Satokata et al., 1992; Satokata et al., 1992). Purified recombinant XPA has been shown to correct the repair defect in extracts from XP-A cells (Park and Sancar, 1993; Robins et al., 1991) clearly demonstrating the defect in these cells.

Damage recognition can occur through the combined actions of XPA and the single-stranded binding protein RPA (Robins et al., 1991) (see Figure 1.2). Individually both preferentially bind single-stranded or damaged DNA, but have a greater affinity as a complex (Burns et al., 1996; He et al., 1995). The same is true for the yeast homologues Rad14 and Rpa (Guzder et al., 1993). It is also interesting that XPA and Rad14 generally bind to (6-4) photoproducts rather than pyrimidine dimers in UV irradiated DNA (Guzder et al., 1993; Jones and Wood, 1993) suggesting that these lesions are likely to be repaired more efficiently by NER. Also, (6-4) photoproducts do not distort the DNA double helix to the same extent as dimers, indicating that XPA does not recognise this type of damage through structural changes in the DNA.

Replication protein A (RPA) is a heterotrimeric complex consisting of three polypeptides of 70, 32 and 14 kDa. It has been shown to be involved in replication (Fairman and Stillman, 1988; Wold and Kelly, 1988), probably during the initiation and elongation stages. RPA has also been shown to be necessary for NER, since depleted extracts are deficient in incision (Shivji et al., 1992). Also, mutants with deletions in XPA that are unable to bind RPA are defective in NER (Li et al., 1995). These and other data have demonstrated that the XPA-RPA interaction is necessary for damage processing (Li et al., 1995) and incision (Mu et al., 1995).

XPA is known to interact with damaged DNA, RPA, TFIIH and ERCC1, and the sites of these interactions have been extensively studied. The XPA protein consists of 273 amino acids and has a molecular weight of 31 kDa. There are three consensus motifs within the gene, two zinc finger motifs and a nuclear localisation signal (Miyamoto et al., 1992; Tanaka et al., 1990). It has been shown that the amino acids from met98 to phe219 are necessary and sufficient for binding to damaged DNA (Kuraoka et al., 1996). In particular, mutation of cys108 to a serine residue abolishes the DNA binding activity of XPA (Asahina et al., 1994; Miyamoto et al., 1992). Binding to ERCC1 occurs via a region containing a cluster of seven glutamic acid residues (Nagai et al., 1995) and interaction with TFIIH is through the C-terminal 48 amino acids (Park et al., 1995). Expression of mutant cDNAs missing either the glutamic acid cluster or the C-terminal 48aas in XP-A cells, results in only marginal recovery of the UV resistant phenotype (Li et al., 1995; Miyamoto et al., 1992). Neither of these regions of the protein is required for DNA binding activity, but interaction with ERCC1 is stimulatory (Nagai et al., 1995).

As mentioned earlier, XPA is able to recognise damage when in a tight complex with RPA and is able to bind to two of the three subunits of RPA. Residues 153 to 176 are required for binding to the RPA 70 kDa subunit and the first 58 residues bind to the RPA 34 kDa subunit (Li et al., 1995). However, it appears that only the interaction with the RPA 70 kDa subunit is significant, as deletion of residues 153-176 of XPA abolishes NER activity and there is no effect from deletion of amino acids 1-58 (Li et al., 1995). It is also interesting that the essential region for interaction with RPA is also necessary for binding to damaged DNA.

More specifically, RPA has been shown *in vitro* to suppress the non-specific endonuclease activity of XPF-ERCC1 whilst promoting its specific junction cutting activity (see below) (Matsunaga et al., 1996). Recently it has also been reported that although RPA may not be required for XPF-ERCC1 activity (see below), it specifically blocks incision of the strand

to which it is bound whilst stimulating cleavage of the other strand (de Laat et al., 1998). A similar activity has been suggested for RPA on XPG incision of an open 30 nucleotide bubble substrate (Matsunaga et al., 1996), although this is not universally agreed (Evans et al., 1997).

Several groups have proposed similar models for the role of RPA in NER. It has been suggested that RPA together with XPA is able to recognise damaged DNA, with RPA able to stabilise the unwound DNA around the lesion. Furthermore it is postulated that RPA plays a role in the recruitment of the nucleases, and the release of the excised oligonucleotide and repair proteins. This model is supported by data indicating that RPA can bind XPG (He et al., 1995) and stimulate the activity of the DNA polymerases α and δ (Coverley et al., 1991). It is also interesting that the predicted length of the repair patch (26-27nt) correlates well with the size of the RPA binding site of 30nt (Kim et al., 1992) suggesting that RPA may be located across the repair patch, either protecting the site from non-specific nucleases, or stabilising the structure.

Knockout mice for XPA have been generated which are completely defective in NER (de Vries et al., 1995; Nakane et al., 1995) confirming the requirement for this protein in repair. The mice are hypersensitive to UVB and chemical carcinogens resulting in skin and eye tumours (de Vries et al., 1995; Nakane et al., 1995). Nakane et. al. reported that the mice are fertile and demonstrate no physical abnormalities or pathological alterations. Particularly relevant for comparison with the XP-A patients is the lack of any neurodegeneration or gliosis in mice studied up to the age of 18 months (Nakane et al., 1995) whereas patients have neurological abnormalities. However mice of over 13 days have been reported as having growth retardation and liver defects such as reduced size, anaemia and disturbed haematopoiesis (de Vries et al., 1995). The XPA knockout mice demonstrate some of the symptoms exhibited by XP-A patients and show the same repair defects and UV sensitivity, thereby providing a good model system.

1.1.4.2 XPC/HHR23A&B

The *XPC* gene product and associated proteins also have some involvement in damage recognition in UV irradiated DNA. Five XP-C patients have been studied to determine the site of the causative mutation (Li et al., 1993). Most of the mutations found in *XPC* were frameshifts and therefore totally inactivating, demonstrating that XPC is not essential for viability.

Much of the work on human XPC has been carried out using mutant human cell lines. XP-C cells are unable to carry out the repair of damaged DNA in transcriptionally inactive regions of the genome, but are able to repair damage on the transcribed strand of active genes (Venema et al., 1991; Venema et al., 1990). This observation explains the residual repair that can be seen in XP-C mutant cells. Using an *in vitro* DNA repair complementation assay, Masutani et al. purified XPC and found that it co-purifies with a protein that is the homologue of yeast Rad23 (Masutani et al., 1994). In addition to this protein termed hHR23B a second homologue of Rad23 exists, called hHR23A which is also able to complex with XPC, although this is barely detectable (Sugasawa et al., 1996). XPC alone is able to bind DNA with a preference for single-stranded DNA (Masutani et al., 1994) or damaged DNA (Reardon et al., 1996). However, when XPC is complexed with either hHR23A or hHR23B its repair activity is stimulated. Generally XPC is found in a complex with hHR23B, although the reason for such a preference is not clear since either Rad23 homologue can produce stimulation of repair. Reconstitution experiments (Aboussekhra et al., 1995; Bessho et al., 1997) have demonstrated that the XPC complex is necessary for NER. This has led to a widely accepted model that XPC/hHR23 is required for the formation or stabilisation of the damage recognition-incision complex. This is further supported by data that showed hHR23 stimulates repair activity at an early stage of the reaction either at or prior to incision (Evans et al., 1997; Sugasawa et al., 1997). There

is also recent direct evidence that XPC is involved in damage recognition (Sugasawa et al., unpublished observations).

The homologue of XPC in yeast, Rad4 forms a complex with Rad23 that has been shown to be essential for reconstitution of *S.cerevisiae* NER (Guzder et al., 1995). Rad4 is not essential for viability, and Rad23 cells show only moderate UV sensitivity. However, Rad23 cells demonstrate no repair of CPDs or (6-4) photoproducts up to 4 hours after UV irradiation (Verhage et al., 1996). The reason for the moderate UV sensitivity is unclear, since it is not consistent with the repair characteristics of the cells.

1.1.4.3 XPG

Cells from XP-G patients demonstrate low levels of excision activity (Reardon et al., 1993) and *in vivo* experiments have shown a severe lack of repair (Cleaver and Kraemer, 1989). Recombinant XPG protein is able to correct the defective NER in human cell extracts from XP-G patients (O'Donovan and Wood, 1993), demonstrating that this protein activity is absent in these cells.

The product of the *XPG* gene and its *S.cerevisiae* homologue *Rad2* are structure-specific endonucleases (Habraken et al., 1995; O'Donovan et al., 1994) which are able to cleave at junctions between ssDNA and dsDNA where the free single-stranded tail is 5' to the double-stranded DNA. It has also been demonstrated *in vitro* that purified recombinant XPG protein can cleave an artificial bubble structure with an opening of at least 5bp (Evans et al., 1997). This group also showed that the cutting could occur on either strand of the structure, but the single-strand/double-strand junction had to have 5' to 3' polarity. These yeast and human proteins are also members of the DnaseIV/FEN-1 gene family, many of which are involved in replication (reviewed in Lieber, 1997). FEN1 is a nuclease which is able to cleave at branched DNA structures and in particular is required for the processing of Okazaki fragments. The FEN1 family of proteins have three regions of homology termed N, I and C. Most family members have only a short gap of 18 amino acids between

regions N and I, illustrated in Figure 1.3, whereas the NER proteins tend to have a gap of over 640 residues.

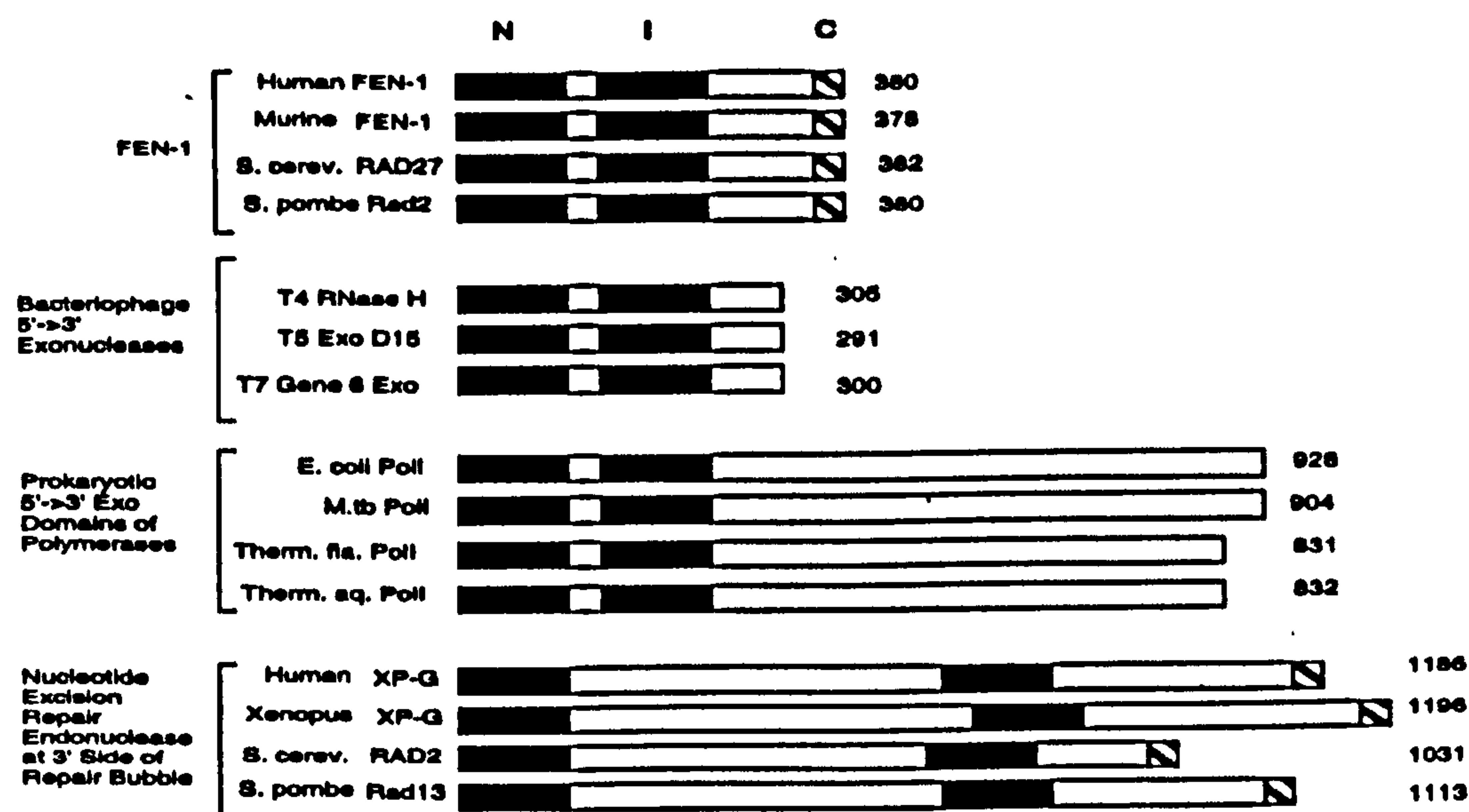


Figure 1.3 The FEN1 family of proteins illustrating the N, I and C regions of homology (Lieber, 1997).

This may be due to the fact that the NER members do not require a free 5' end for nuclease activity, suggesting there may be structural reasons for the larger separation of these regions. Also, it is interesting that the N and I spacer regions of Rad2 can be removed and result in a nuclease capable of flap cleavage, supporting the idea that the larger spacer region is significant for the cleavage of NER substrates. In summary, this group of proteins are nucleases and all tend to produce nicks at branched DNA structures, providing further evidence for the involvement of XPG in the 3' incision of NER. Consistent with these findings, cells deficient in XPG are unable to make the 3' incision in an artificial substrate containing a single cisplatin lesion, and subsequently are defective in repair synthesis and excision of the damaged oligonucleotide (O'Donovan et al., 1994).

1.1.4.4 ERCC1-XPF

Cells from most patients in XP complementation group F are moderately UV sensitive (Cleaver and Kraemer, 1989) and extracts from these cells are defective in excision

nuclease activity (Reardon et al., 1993). When treated with UV, XP-F cells are able to repair 10-15% of damage shortly after irradiation, increasing to 60% after longer incubation (Cleaver and Kraemer, 1989).

The proteins XPF and ERCC1 form a tight complex, formation of which is intrinsic to the stability of both proteins. Mutations occurring in the middle of *ERCC1* result in an unstable protein, probably due to the loss of the XPF interaction (Sijbers et al., 1996). As with XPG, the ERCC1-XPF complex and its yeast counterpart Rad1-Rad10 is able to cleave at single-strand/double-strand junctions. The polarity of the junction cleaved is the opposite of that observed for XPG suggesting that ERCC1-XPF is responsible for the second incision in NER, 5' of the damage site (Matsunaga et al., 1995) (see figure 1.2).

A great deal of work has been carried out on the Rad1 and Rad10 proteins, the *S.cerevisiae* homologues of XPF and ERCC1. This has provided insight into the mechanism of action of the human complex in NER and its involvement in other pathways. Deletion mutants of Rad1 and Rad10 are very UV sensitive and exhibit no damage-specific DNA incision. The Rad1-Rad10 complex also functions in the mitotic recombination pathway implicating ERCC1-XPF in recombination. This might explain the sensitivity of *ERCC1* and *XPF* mutants to crosslinking agents, as resulting lesions require mitotic recombination for their repair (Schiestl and Prakash, 1990).

Knockout mice have been generated for *ERCC1* and have been shown to have disturbed growth, a reduced lifespan and abnormalities of liver nuclei (McWhir et al., 1993). Weeda et. al. have also produced mice null for *ERCC1*, and mice carrying a seven amino acid C-terminal truncation that results in a transcript reduced to approximately 15% of that seen in normal mice (Weeda et al., 1997). The symptoms displayed by *ERCC1* mice are very severe compared to those exhibited by *XPA* knockout mice, since both mutations result in totally defective NER. *XPA* mice are viable, develop normally and have a normal lifespan (de Vries et al., 1995; Nakane et al., 1995). This suggests an additional function for

ERCC1, as alluded to above, probably in the repair of genotoxic lesions such as crosslinks, via the mitotic recombination pathway.

The location and coupling of incisions in NER has been extensively studied. Generally, the 3' incision mediated by XPG occurs first, since uncoupled 3' incisions are found more commonly than uncoupled 5' incisions (Mu et al., 1996; O'Donovan et al., 1994; Sijbers et al., 1996). In some cases and with certain types of lesion, it has been demonstrated that the 3' incision can occur in the absence of ERCC1-XPF (Mu et al., 1996), whereas, for other damages such as cisplatin, absence of the complex results in the complete loss of incision activity (Moggs et al., 1996). Interestingly, in the yeast system it has been shown that the Rad1-Rad10 complex is indispensable for incision with the reconstituted system (Guzder et al., 1995).

The location of the sites of incision correlates well with that of *E.coli*, with the 3' incision occurring closer to the lesion. It has been shown that the 3' nick occurs between two and nine phosphodiester bonds from the site of damage. The second site of incision generally occurs sixteen to twenty-five bonds 5' of the damage, releasing an oligonucleotide of around twenty-six to twenty-seven nucleotides in length (Moggs et al., 1996). The precise location of the incisions is dependent on the nature of the lesion, but the length of the displaced oligonucleotide appears to be relatively consistent in all cases.

1.1.4.5 TFIIH

TFIIH is a multi-subunit complex containing between six and nine subunits depending on the preparation and was first identified as one of the RNA polymerase II general transcription factors. TFIIH was identified as being involved in repair through the discovery that the proteins encoded by the genes *XPD* and *XPB*, defective in patients with the repair disorder xeroderma pigmentosum, were part of the complex (Drapkin et al., 1994; Schaeffer et al., 1994; Schaeffer et al., 1993). TFIIH was subsequently shown to

complement the repair defects in cells from XP patients in complementation groups B and D (Mu et al., 1995; Schaeffer et al., 1994; Schaeffer et al., 1993). Although the intact complex is required for the NER reaction to occur, it is the activities associated with the XPB and XPD proteins that are thought to be necessary to effect repair. These two proteins are ATP-dependent DNA helicases with XPB and XPD demonstrating 3' to 5' and 5' to 3' polarity respectively (Drapkin et al., 1994; Schaeffer et al., 1993).

Patients from XP complementation groups B and D show some similarities and the proteins XPB and XPD are helicases and part of the TFIIH complex (Drapkin et al., 1994; Schaeffer et al., 1994; Schaeffer et al., 1993). Both complementation groups contain patients with symptoms of TTD, and XP with CS, and to date no patient from complementation group B has exhibited the clinical features of XP alone (Broughton et al., 1994; Vermeulen et al., 1994; Weeda et al., 1990). At the cellular level XP-B cells have a weak excision signal and XP-D cells fail to excise dimers at a detectable level as measured in an *in vitro* excision assay (Reardon et al., 1993), demonstrating a likely requirement for both proteins for active NER *in vivo*.

There are only four patients from three families assigned to the XP-B group and three of these patients have features of XP with CS. The repair defect in two siblings can be corrected by microinjection of *XPB* cDNA (Vermeulen et al., 1993). The repair defect in XP-B cells can also be alleviated by microinjection of the TFIIH complex (van Vuuren et al., 1994). Mutational analysis of the *XPB* gene has shown that the helicase domains are essential for the repair function and that the protein is nuclear, independent of its own nuclear localisation signal (Ma et al., 1994).

Many patients have been assigned to the XP-D complementation group. Mutations in the *XPD* gene inactivate its ability to correct the UV sensitive phenotype of XP-D cells (Frederick et al., 1994) and can result in either XP, XP with CS or TTD (Broughton et al., 1994; Broughton et al., 1995; Takayama et al., 1996; Takayama et al., 1995). *XPD* is an

essential gene and studies on the *S.pombe* and *S.cerevisiae* yeast homologs *rad15* and *RAD3* have been used to study mutations. By identifying the totally inactivating mutations, Taylor et. al. were then able to identify mutations characteristic to either of the disease phenotypes XP, XP/CS or TTD (Taylor et al., 1997). The *S.cerevisiae* homologue Rad3 has been used to investigate the significance of the XPD helicase activity (Guzder et al., 1995). The data from both yeast systems suggests that the XPD helicase activity is required for NER but is not essential for transcription, and they are consistent with the idea that TTD is a transcription disorder.

1.1.4.6 TTD-A

The repair defect in TTD-A cells can be corrected by microinjection of the RNA polymerase II general transcription factor TFIIH (Vermeulen et al., 1994) (see below), suggesting that the gene mutated in these cells is a component of this complex. However, the major genes which form TFIIH have been cloned and none has been shown to correct the repair defect of TTD-A (Marinoni et al., 1997; Vermeulen et al., 1994).

1.1.4.7 XPE

In 1988 it was reported that an XP-E cell line lacked a factor which could bind DNA (Chu and Chang, 1988; Patterson and Chu, 1989). However, further work has shown that only three from twelve XP-E patients studied are defective in this DNA binding activity (Keeney et al., 1994). The DNA binding protein is a heterodimer of proteins of molecular weights 127 kDa and 48 kDa (Keeney et al., 1992). Investigation into the genes encoding these proteins has shown that mutations found to date are only seen in the smaller p48 subunit (Nichols et al., 1996).

1.1.5 Repair synthesis

The repair synthesis step of NER in mammalian cells has been much less studied than the incision step, but it is known that the post-incision repair synthesis reaction requires a

minimum of four proteins; PCNA, RFC, DNA ligase I and DNA polymerase δ or ϵ . Reconstitution of the NER reaction has shown these proteins to be involved, with pol ϵ the polymerase in this case (Aboussekhra et al., 1995). It is thought that RFC loads PCNA onto a repair-type template for synthesis by either DNA pol δ or ϵ (Podust et al., 1994). There is considerable evidence for the involvement of PCNA in repair synthesis, both *in vitro* (Nichols and Sancar, 1992; Shivji et al., 1992) and *in vivo* (Celis and Madsen, 1986). Depletion of PCNA from human cell extracts results in a complete lack of repair synthesis (Shivji et al., 1992).

There is a single patient, 46BR that possesses a mutation in the DNA ligase I gene and shows a defect in the joining of Okazaki fragments during DNA replication. This patient is also hypersensitive to DNA damaging agents including UV, suggesting that, along with the reconstitution data, ligase I is responsible for sealing the nick to the newly synthesised oligonucleotide (Barnes et al., 1992). The excised oligonucleotide is released independently of repair synthesis and both the excised oligo and the gapped DNA are complexed with proteins (Mu et al., 1996), although the exact proteins bound is not clear. It has also been shown that the kinetics of the incision and repair synthesis reactions are similar, hinting that these two steps are likely to be tightly coupled (Moggs et al., 1996).

1.2 Basal Transcription by RNA Polymerase II

There is a clear link between nucleotide excision repair and transcription, through the transcription/repair factor TFIIH, as discussed above in a repair context. This section will explain the basics of transcription initiation and explain in some detail the role that TFIIH fulfils in this process. For reviews on transcription initiation (see Buratowski, 1994; Maldonado et al., 1996; Nikolov and Burley, 1997; Ranish and Hahn, 1996; Stargell and Struhl, 1996; Tjian and Maniatis, 1994; Zawel et al., 1995).

Transcription in human cells is a complex process with many associated factors as well as those which are directly involved in initiation. As generally understood, TFIID binds to the TATA-box of the promoter where the two together act as a binding site for TFIIB which can recruit RNA polymerase II (RNAPII) and TFIIF. This complex is kinetically stable and is able to recruit TFIIE and TFIIH, creating an ATP-dependent open complex, thereby allowing transcription to initiate.

The minimal transcription initiation complex consists of TFIID, TFIIA, TFIIB, TFIIF and RNA pol II. TFIID is a complex comprising the TATA-box binding protein (TBP) and combinations of associated factors (TAFs) of which there are eight known (reviewed in (Tansey and Herr, 1997)). TBP alone is sufficient for binding to the TATA-box, but binding to other core promoter elements is dependent on the TAFs. TBP itself has a saddle shape structure of which the inner surface interacts with the minor groove of the TATA-box causing a distortion in the DNA (Nikolov and Burley, 1997).

TFIIA binds the promoter-TFIID complex, but is not required for promoters where transcription is driven by TBP alone. TFIIA is thought to work by establishing and maintaining the complex under physiological conditions by counteracting repressors of transcription.

TFIIB has two domains and binds the TBP-DNA complex where it is then able to recruit RNA polymerase II. The protein is able to interact *in vitro* with RNAPII, TFIIF and the TBP-DNA complex. RNAPII is recruited at this stage, followed by TFIIF via interactions with RNAPII and TFIIB. TFIIF unlike the other initiation factors is also thought to be involved in the elongation stages of transcription.

The complete initiation complex consists of the factors described above along with TFIIE and TFIIH. However, it is possible for transcription to be initiated *in vitro* in the absence of these two factors. TFIIE is a tetramer containing two subunits of each of the products of

two genes. TFIIE is required for the recruitment of TFIIH and appears to have a regulatory role in the enzymatic functions of this complex.

There are two theories about the mechanism of basal transcription, whether the transcription factors and associated proteins exist as a complex, termed the holoenzyme, or are recruited in a step-wise fashion, as described above. A holoenzyme was first described in yeast and required the addition of TBP and TFIIE for transcriptional activity (Koleske and Young, 1994). However, the RNAPII holoenzyme described by Ossipow et. al., 1995 from human cell extracts contains both these factors. Since holoenzymes have been isolated from human and yeast cells as an intact complex, it seems likely that they do exist at some stage during transcription. As to whether such large complexes are consistently present in the cell remains a matter for some debate.

In addition to the basal transcription factors described above, several proteins have been identified as associating with the RNAPII holoenzyme. Such factors include the group known as suppressors of RNA polymerase B (SRBs) which were first identified in yeast (reviewed in (Koleske and Young, 1995)). This group of proteins is thought to be involved in stabilising interactions between RNAPII and the general transcription factors and may also confer responsiveness to co-activators (Koleske and Young, 1994). The human equivalents of SRBs have been identified both by analysis of the constituents of a purified RNAPII complex (Maldonado et al., 1996) and by immunoprecipitation with anti-SRB antibodies, which also bring down RNAPII (Chao et al., 1996).

1.2.1 TFIIH

As previously discussed, the general transcription factor TFIIH is involved in both transcription and NER. TFIIH is a multi-subunit complex and consists of nine polypeptides; XPB, XPD, Cdk7/MO15, cyclin H, MAT1, p34, p44, p52 and p62 (reviewed in Hoeijmakers et al., 1996). No precise functions have as yet been assigned to the p34, p44, p52 and p62 subunits, however all except p62 have been shown to contain consensus

sequences. Zinc finger motifs often occurring in DNA binding proteins are present in p34 and p44 and p52 has WD-repeats which are thought to have a structural role possibly in protein:protein interactions, indicating possible roles for these proteins. The subunits XPB and XPD as explained above are helicases that work with opposite polarity and are likely to be involved in open complex formation, allowing the transcription and repair machinery access. The final three subunits are involved in the kinase activity associated with TFIIH.

The C-terminal domain of the large subunit of RNA polymerase II (RNAPII-CTD) is the target of phosphorylation by TFIIH (Serizawa et al., 1995). This activity is conferred by the cyclin-dependent activating kinase (CAK), which consists of the TFIIH subunits Cdk7/MO15 and cyclin H (Makela et al., 1994). CAK can exist either in complex with TFIIH or in a smaller complex (Adamczewski et al., 1996; Shiekhattar et al., 1995). *In vitro* studies have demonstrated that CAK can either function alone, or as part of TFIIH to phosphorylate RNAPII-CTD, TFIIE, TFIIIF, TBP and the cell cycle regulators cdc2 and cdk2 (Okhuma and Roeder, 1994; Serizawa et al., 1995; Shiekhattar et al., 1995).

Phosphorylation of the RNAPII-CTD is an important event, since its phosphorylation state shows a distinct correlation with the transcriptional status of the enzyme. Un-phosphorylated RNAPII is required for initiation (Akoulitchiev et al., 1995) and resides in complexes that are ready to initiate. However, the RNAPII that exists during elongation is highly phosphorylated (Cadena and Dahmus, 1987; O'Brien et al., 1994).

In vitro studies have shown CAK activity associated with TFIIH to have a specificity for RNAPII-CTD, whereas, free or recombinant CAK demonstrates a preference for cdk2 and cdc2 (Rossignol et al., 1997; Yankulov and Bentley, 1997). Shiekhattar et. al. have also reported specificity by showing that although both complexes are able to phosphorylate RNAPII-CTD, only TFIIH can carry out the reaction when RNAPII is part of the pre-initiation complex (Shiekhattar et al., 1995).

MAT1, another TFIIH subunit is also involved in the kinase activity associated with this transcription factor and its substrate specificity. MAT1 strongly associates with cyclin H and cdk7, but is not essential for CAK assembly or kinase activity (Adamczewski et al., 1996). MAT1 does, however have a stimulatory effect on kinase activity (Adamczewski et al., 1996) and alters the substrate specificity of CAK from cdk2 to RNAPII-CTD (Yankulov and Bentley, 1997). In addition to MAT1, TFIIH is also able to stimulate the RNAPII-CTD phosphorylation by TFIIH (Lu et al., 1992). This is conferred by p56, one of two TFIIH subunits (Okhuma and Roeder, 1994) which is reported to occur in an oligomeric form for this purpose, distinct from the transcriptionally active heterodimer (Serizawa et al., 1994). It can be postulated from the above data that phosphorylation of RNAPII-CTD performs a regulatory role in transcription and possibly traps or stabilises RNAPII in the elongation complex.

TFIIH provides a link between the cellular processes of transcription, NER and cell cycle regulation. It is likely that the CAK activity which phosphorylates cdc2 and cdk2 *in vivo* is separate from that of TFIIH, since free CAK rather than TFIIH has the greatest affinity for these substrates. Also, it has been demonstrated that TFIIH kinase activity remains constant during the cell cycle (Adamczewski et al., 1996).

The link with NER is somewhat clearer, with mutations in the TFIIH subunits XPB and XPD resulting in the DNA repair disorders XP, TTD and XP with CS. In addition to this, Roy et. al. have shown that MO15/cdk7 depleted extracts could not function in an *in vitro* repair assay (Roy et al., 1994). This is likely to be a result of destabilisation or inactivation of the whole TFIIH complex, rather than the lack of kinase activity. It is also important to consider the observations of Adamczewski et. al., 1996 who reported that TFIIH kinase activity is reduced after UV irradiation. This again may be a secondary response rather than a direct effect from the kinase, but DNA damage clearly affects the activity. This may

be through recruitment of TFIIH to the damage site for repair, during which the kinase activity is unnecessary.

1.3 Transcription-coupled repair

Transcription and repair are linked through the basal transcription and NER factor TFIIH. A previous link had been established when it was shown that transcription inhibited by UV irradiation recovered faster than could be accounted for by the rate of damage removal from the genome (Mayne and Lehmann, 1982). This suggested another more rapid mechanism of repair was responsible for recovery from inhibition of transcription. Such a mechanism was demonstrated by Hanawalt and co-workers in 1985, who reported a five-fold faster repair of UV induced pyrimidine dimers in the transcriptionally active *DHFR* gene in CHO cells when compared with the genome overall (Bohr et al., 1985). A similar phenomenon was reported in mouse 3T3 cells with the active *c-abl* gene, which was repaired much faster than the inactive *c-mos* gene (Madhani et al., 1986).

More detailed information on “transcription-coupled repair” (TCR) was provided by Mellon et. al. who showed that preferential repair was due to faster removal of CPDs from the transcribed strand. The non-transcribed strand was repaired at the same rate as the rest of the genome (Mellon et al., 1987). There are two cellular factors, chromatin structure and transcription, which may influence or control transcription-coupled repair.

It has been proposed that repair of the transcribed strand is initiated by a stalled RNAPII, blocked at a lesion (Hanawalt et al., 1994). Also, treatment with the RNA polymerase inhibitor α -amanitin abolishes selective repair of the transcribed strand (Leadon and Lawrence, 1991), demonstrating a requirement for active transcription in TCR. However, the rate of TCR does not correlate with that of transcription (Bedoyan et al., 1992; Venema et al., 1992) or with the transcription-blocking efficiency of the lesion (McGregor et al., 1995). Therefore, a further process must be functioning to effect preferential repair.

It has been reported that the compactness of the local chromatin structure around a lesion influences the rate of its repair (Smerdon and Thoma, 1990). Additionally, cells treated with sodium butyrate, causing relaxation of nucleosomal structure, show enhanced excision repair (Ramanathan and Smerdon, 1989). Such data implies that transcription-coupled repair is not a selective mechanism, but occurs purely as a result of increased accessibility of the repair machinery to the damaged DNA. However, the existence of transcription-coupled repair in the *lac* operon of *E.coli* (Mellon and Hanawalt, 1989) casts doubt on this theory, since *E.coli* does not possess a stable nucleosome structure. Therefore, it is unlikely that an open chromatin structure around sites of damage directly results in TCR, although it clearly influences the general rate of repair.

1.3.1 Transcription-coupled repair in *E.coli*

Preferential repair in *E.coli* was first noted by study of the *lac* operon *in vivo* under induced and un-induced conditions, that is, with active and inactive transcription respectively. Mellon and Hanawalt reported that when the *lac* operon was induced, the transcribed strand was repaired faster than the complementary strand (Mellon and Hanawalt, 1989). The existence of such strand-specific repair was not entirely new, since the phenomenon of mutation frequency decline (MFD) had alluded to such a process many years earlier. Mutation frequency decline can be explained such that after UV irradiation, cells are able to rapidly decrease the frequency of revertant mutations in tRNA genes, without loss of viability (reviewed in Witkin, 1994). Bockrath and Palmer reported that the revertant phenotype occurred via UV induced mutations in the transcribed strand of tRNA genes. Cells mutant for MFD, *mfd-1* are unable to prevent these mutations, and are therefore predicted to have a defect in transcription-coupled repair (Bockrath and Palmer, 1977). The *mfd-1* mutant cells have since been shown to possess a defect in TCR both *in vitro* (Selby et al., 1991) and *in vivo* (Kunala and Brash, 1995). Furthermore, a partial MFD

phenotype could be restored in *mfd-1* cells by inhibiting transcription with rifampicin, demonstrating a requirement for active transcription in this process (Li and Bockrath, 1995).

The transcription-coupled repair reaction in *E.coli* could not be reconstituted with known repair proteins (Selby and Sancar, 1990), suggesting that the presence of a stalled RNA polymerase molecule was not sufficient to induce preferential repair. More detailed studies have shown that pyrimidine dimers in the non-template strand have no effect on transcription, with no stalling of RNAP. However, dimers occurring in the template strand block RNAP elongation and cause stable ternary complexes (Selby and Sancar, 1990). Therefore, although a stalled RNAP molecule is not the direct signal for repair, its presence appears to be essential for TCR to occur. *In vivo* studies on the actively transcribed genes *uvrC* (Selby and Sancar, 1995), *lacI* and *lacZ* (Kunala and Brash, 1995) have shown that TCR occurs only with an elongating RNAP complex. This indicates that there may be an inhibition of preferential repair when RNAP is bound to the initiation site. Therefore, a “transcription-repair coupling factor”(TRCF) is presumed to exist to allow repair proteins to recognise RNAP stalled at a lesion, rather than within the initiation complex.

The transcription-repair coupling factor is identical to the *MFD* gene product, and the TCR defect observed in *mfd-1* cells can be corrected by addition of purified TRCF (Selby and Sancar, 1993; Selby et al., 1991). TRCF is an abundant protein of 130 kDa that shares homology with UvrB at its N-terminus as well as a C-terminal potential leucine zipper and seven recG-like helicase motifs (Selby and Sancar, 1993). As yet no helicase activity has been demonstrated for TRCF (Selby and Sancar, 1995), although it does have a weak ATPase activity that is not stimulated by DNA or stalled RNAP (Selby and Sancar, 1993). The UvrB-like domain may facilitate binding to UvrA (Selby and Sancar, 1993) and provide a means for recruitment of the repair proteins to the lesion.

Using a defined reconstituted system, Selby and Sancar demonstrated that transcription preferentially inhibited repair of the transcribed strand. Such inhibition could be overcome by addition of TRCF. Not only was the inhibition prevented but, repair of the transcribed strand was stimulated by TRCF without any effect on repair of the non-transcribed strand (Selby and Sancar, 1993). TRCF is able to bind DNA at initiation and elongating RNAP complexes in an ATP-dependent manner (Selby and Sancar, 1995), but TCR only occurs with an elongating RNAP complex (Kunala and Brash, 1995; Selby and Sancar, 1995). Therefore, it is not the binding of TRCF to the stalled complex, but its ability to displace it from the site of the lesion, which facilitates the repair. On this pretext, TRCF has no effect on an initiation complex but is able to release RNAP in an ATP-dependent manner from an elongation complex stalled at a lesion without affecting initiation complexes. The mechanism of displacement is, however distinct from that of the rho-dependent termination of transcription (Pavco and Steege, 1990; Selby and Sancar, 1995) discounting any involvement of this process in TCR. Displacement of stalled RNAP by TRCF is not specific to complexes stalled at DNA damage, since the same process occurs when elongation is blocked by a tightly bound protein (Pavco and Steege, 1990).

A model can be proposed for the mechanism of transcription-coupled repair in *E.coli*. TRCF binds at a stalled RNAP complex and displaces the polymerase and the truncated transcript. The ability of TRCF to bind UvrA targets this protein to the damage site, allowing recruitment of UvrB and the rest of the repair enzymes.

Unfortunately, Selby and Sancar were unable to detect targeting of (A)BC exonuclease to the lesion in their reconstituted system (Selby and Sancar, 1993). However, binding of TRCF to UvrA is significant, since addition of excess UvrA inhibits TCR, presumably by binding to TRCF away from the DNA and preventing its action on stalled RNAP complexes (Selby and Sancar, 1995).

1.3.2 Transcription-coupled repair in *S.cerevisiae*

The first evidence for preferential repair in yeast was provided by study of the mating type loci *MAT α* and *HML α* . The two loci have identical sequences, but in yeast α -cells the *MAT α* locus is actively transcribed, whilst the *HML α* locus is silenced by the Sir proteins. Terleth et. al., 1989 reported that *MAT α* is repaired at two and a half times the rate of *HML α* . However, mutation of the upstream activating sequence of *MAT α* did not eliminate the preferential repair, despite the lack of active transcription. This preferential repair was absent in a strain mutant for Sir3, one of the proteins responsible for silencing of *HML α* (Terleth et al., 1989). This was due to an increase in the repair of *HML α* rather than a reduction for *MAT α* , suggesting that the local chromatin structure was responsible for the differences in repair rates.

Authentic “transcription-coupled repair” or transcription-dependent repair has subsequently been demonstrated in *URA3*, *RPB2* and *GAL7*, where there is a distinct requirement for active transcription. The *GAL7* gene is inducible, thereby allowing a direct comparison of repair within a gene under conditions of active and inactive transcription. When induced, the transcribed strand is repaired two to three times faster than the non-transcribed strand or the overall genome. In contrast, there is no preferential or strand-specific repair observed when *GAL7* is uninduced, with both strands repaired at the same rate as the rest of the genome (Leadon and Lawrence, 1992). A similar observation was made in the actively transcribed *RPB2* gene, the TCR of which was shown to be dependent on transcription. Strand-specific repair could be observed when the gene was actively transcribed either within the genome or episomally on a plasmid (Sweder and Hanawalt, 1992).

1.3.2.1 Yeast TCR mutants

The gene *RAD26* was identified through its homology with the human *CSB* gene that is defective in patients with Cockayne syndrome. Cells derived from these patients lack transcription-coupled repair (van Gool et al., 1994). Deletion mutants of *rad26* are viable and grow normally demonstrating that the protein does not have an essential function (Verhage et al., 1996). The cells also do not show any sensitivity to UV, cisplatin or γ -irradiation, although they do exhibit a delay in recovery of growth after UV treatment (van Gool et al., 1994). *rad26* mutants are almost completely deficient in the preferential repair of the transcribed strand of *RPB2* (van Gool et al., 1994), being repaired at virtually the same rate as the non-transcribed strand. Unlike *E.coli* *mfd-1* there is no evidence of inhibition of repair in the transcribed strand by a stalled RNA polymerase II molecule. Purified Rad26, like MFD is an ATPase, but the enzymatic activity is far more active and has a strict DNA dependence (Guzder et al., 1996).

The yeast homologue of the second Cockayne syndrome gene, *CSA* has also been cloned and has been called *RAD28* (Bhatia et al., 1996). In contrast to *rad26* deletion strains, *rad28* deficient cells are able to carry out transcription-coupled repair at the same rate as wild-type cells (Bhatia et al., 1996). As seen in *rad26*, *rad28* shows no increased sensitivity to killing by UV, γ or MMS (Bhatia et al., 1996), thus indicating that Rad28 is not required for strand-specific repair of UV induced damage, contrary to the case for *CSA* in human cells (see below).

1.3.2.2 Double mutants

Double mutants have been created for the TCR genes *rad26* and *rad28* with the global genome repair genes *rad7* and *rad16* which are essential for the repair of non-transcribed regions of the genome (Bang et al., 1992; Verhage et al., 1994). Double mutants of *rad28* with either *rad7* or *rad16* show only a slightly enhanced UV sensitivity compared to *rad28* alone (Bhatia et al., 1996). This illustrates that even in a global repair deficient

background, Rad28 does not contribute to UV sensitivity. Such evidence suggests that either *RAD28* is not the true functional homologue of *CSA*, or that this gene has no function in yeast.

rad26 deletion strains are impaired for transcription-coupled repair, but are not UV sensitive compared to wild-type (van Gool et al., 1994). However, double mutants of *rad26* and *rad7* or *rad16* are more UV sensitive than either deletion strain alone, but are still less sensitive than mutants which are completely defective in NER (Verhage et al., 1996).

Analysis of strand-specific repair of the *RPB2* locus in *rad7rad26* or *rad16rad26* double mutants showed that there was no repair of the non-transcribed strand and little repair in the transcribed strand. Repair of the transcribed strand in double mutants was less reduced compared to a *rad26* alone background, indicating that *RAD7* and *RAD16* contribute to the repair of the transcribed strand in TCR-defective circumstances (Verhage et al., 1996). Such residual repair of the transcribed strand was also seen in *GAL7* when induced, but not when repressed indicating a dependency on transcription. Therefore, in yeast some transcription-coupled repair can occur independently of Rad26 function. Furthermore, at an induced *GAL7* locus, disruption of *RAD26* has very little effect on TCR unless it is combined with disruption of *RAD7* or *RAD16* (Verhage et al., 1996). Therefore in this situation a coupling factor does not appear to be required, possibly because induced *GAL7* is very heavily transcribed (St. John et al., 1981). This is very similar to the situation in the IPTG induced *lacZ* gene in *E.coli*, where transcription-coupled repair still occurs in a *mfd* mutant (Kunala and Brash, 1995).

The above information indicates that although Rad26 is involved in TCR, it is not essential for all transcribed strand repair. This may be specific to the yeast system, either as a result of redundancy or because a more efficient global genome repair system in yeast reduces the dependency on such a process.

Definition of true transcription-coupled repair suggests that the phenomenon depends on active transcription, thus separating the mechanism from the preferential repair observed through differences in local chromatin structure. A total dependence on RNAPII transcription can be shown by the absence of TCR in temperature sensitive RNAPII mutants at the restrictive temperature (Leadon and Lawrence, 1992; Sweder and Hanawalt, 1992). There is however clearly a link between TCR and chromatin structure, since active transcription will result in a relaxed nucleosome structure in that region.

Recent experiments have involved the study of TCR, the influence of a stalled transcription complex and the recovery from this state. Reagan and Friedberg have developed an assay whereby they could measure the kinetics of recovery of RNA synthesis by analysis of templates carrying at least one RNAPII blocking lesion. This reflects the capacity of cell lines to remove lesions from the transcribed strand of the *GAL10* and *RNR3* genes (Reagan and Friedberg, 1997). The recovery of RNA synthesis in NER mutants directly reflected their repair status, with the NER defective *rad1* and *rad2* mutant cells showing no recovery of RNA synthesis. Also, *rad26* null cells showed a delay in recovery reflecting the ability of global genome repair to repair these lesions, albeit at a slower rate. Interestingly the GGR mutants *rad7* and *rad16* demonstrated a slightly more rapid recovery than wild-type cells, illustrating the lack of involvement of GGR in transcription-coupled repair.

Transcription at the *CYC1* promoter *in vitro* with RNAPII is inhibited by addition of damaged plasmid to the extracts. This inhibition is strictly dependent on NER (You et al., 1998), and requires the *RAD26* gene product. Alleviation of transcription inhibition can occur by addition of TFIIH, with a requirement for the kinase subunits. These data interlink Rad26, TFIIH and stalling of RNA polymerase, suggesting a possible role of Rad26 in this early stage prior to repair.

Further data on the site and timing of action of Rad26 has been reported by Tjisterman et. al., 1997. This group have shown that Rad26 is not required for the repair of lesions on the

transcribed strand within 50bp of the transcription initiation site (Tijsterman et al., 1997). This 50bp distance also corresponds to the region where TFIIH dissociates from the transcription complex. Therefore it has been suggested that since TFIIH is already associated with the transcription machinery within the first 50bp of mRNA synthesis, there is no requirement for a coupling factor (Tijsterman et al., 1997).

1.3.3 Mammalian transcription-coupled repair and Cockayne syndrome

Cockayne syndrome (CS) is a rare autosomal recessive disorder which demonstrates biochemical and genetic heterogeneity. Nance and Berry carried out a comprehensive study of 140 CS patients (Nance and Berry, 1992) demonstrating the clear diagnostic criteria of poor growth, neurological abnormalities and photosensitivity. Other common features of CS include sensorineural hearing loss, cataracts, pigmentary retinopathy and dental caries. A mean age of death of 12¼ years was reported, although more recently two siblings of 42 and 55 years of age have been reported with particularly mild clinical symptoms, but characteristic cellular features (Miyauchi et al., 1994). Unlike patients with xeroderma pigmentosum, no cases of cancer have been reported in classical CS patients and they also have no disposition to infectious complications.

At the cellular level, fibroblasts derived from CS patients are hypersensitive to UV irradiation (Mayne et al., 1982) and fail to recover RNA synthesis after irradiation (Mayne and Lehmann, 1982). This cellular characteristic has been used as a tool for diagnosis of patients and a test for complementation. Complementation analysis has shown there to be two complementation groups of CS alone, A and B, with the majority of patients belonging to group B (Lehmann, 1982; Stefanini et al., 1996; Tanaka et al., 1981). With reference to complementation, no obvious clinical, cellular or racial distinctions can be made between individuals in the two groups (Stefanini et al., 1996).

Cockayne syndrome cells are defective in the repair of transcriptionally active DNA when compared to normal cells (Venema et al., 1990). Study of the two active loci *ADA* and *DHFR* and the inactive locus *754* has demonstrated the transcription-coupled repair defect. In cells derived from either complementation group there is virtually no removal of cyclobutane pyrimidine dimers from either *ADA* or *DHFR* four hours after irradiation. After 24 hrs, the extent and rate of repair is the same in all three loci, that is the same level as for the *754* locus in normal cells (Venema et al., 1990). Interestingly, CS cells also have reduced levels of repair of ribosomal RNA genes which are normally transcribed by RNA polymerase I (Christians and Hanawalt, 1994). This suggests that the CS gene products are required for transcription-coupled repair in a much wider context than just by association with RNA polymerase II.

Neither CSA or CSB protein is essential for viability or basal transcription, however they are required for transcription-coupled repair and recovery of RNA synthesis after UV *in vivo* (van Gool et al., 1997). Analysis of the repair of different DNA adducts has been used to determine the dependence on and influence of transcription. Repair of the NA-AAF induced N-(deoxyguanosin-8-yl)-2-aminofluorene (dG-C8-AF) adduct was studied in the active *ADA* gene, and its repair was shown to occur at similar rates in both normal and CS cell lines without strand specificity (van Oosterwijk et al., 1996). Therefore repair of this lesion is not influenced by transcription-coupled repair. Nevertheless CS cells are three times as sensitive to the toxic effects of NA-AAF and RNA synthesis fails to recover after NA-AAF treatment (van Oosterwijk et al., 1996). Similarly, McGregor et. al. have attempted to correlate the rate of repair in the HPRT gene with the transcription blocking effect of the adducts 1-nitrosopyrene (1-NOP), N-acetoxy-2-acetylaminofluorene (NA-AAF) and benzo(a)pyrene diol epoxide (BPDE). BPDE blocks transcription four times more efficiently than either 1-NOP or NA-AAF. However, the TCR of 1-NOP adducts is twice as fast as that of BPDE lesions, with repair of NA-AAF lesions slower again and not

strand-specific (McGregor et al., 1995). Therefore, the extent to which DNA lesions block transcription has no bearing on the occurrence or rate of transcription-coupled repair.

Donahue et. al. studied the effects of the NA-AAF induced adducts N-2-aminofluorene (AF) and N-2-acetylaminofluorene (AAF) in an *in vitro* transcription assay. AF adducts are not preferentially repaired and when present on the transcribed strand they result in a weak RNAPII pause site that can be bypassed. In contrast, AAF lesions are helix distorting and create an absolute block to elongation when situated on the transcribed strand. Interestingly, either lesion occurring on the non-transcribed strand results in enhanced arrest at nearby sequence-specific pause sites (Donahue et al., 1996). All the above data show that the transcription blocking lesions are also those that are subject to transcription-coupled repair and although there is no direct rate correlation, there is a clear link. It is possible that a stalled RNA polymerase molecule may be involved in the targeting of repair machinery to transcriptionally active DNA, albeit indirectly.

The *CSA* and *CSB* gene products are not only required for TCR of UV damage, but also oxidative damage (e.g. thymine glycols). Cockayne syndrome cells from either complementation group are defective in the repair of ionizing radiation and oxidative damage from the transcribed strand (Cooper et al., 1997; Leadon and Cooper, 1993). Cells derived from patients with XP-G combined with CS are defective in TCR of thymine glycols (Cooper et al., 1997). This demonstrates a correlation between the presence of CS symptoms and the lack of TCR of oxidative and ionising radiation-induced damage. The XPG function required for this TCR role is distinct from its incision role in NER, since cells from XP-G patients with XP alone and totally NER deficient XP-A cells are proficient in this repair. Cooper et. al. have suggested that since XPG has been implicated in the assembly of a pre-incision complex in NER, it may be involved in a similar process in BER, the major pathway for the removal of oxidative damage (Cooper et al., 1997).

1.3.3.1 CSA

The *CSA* gene was cloned by functional complementation of a CS-A cell line with a cDNA library and confirmed by identification of mutations in the gene in CS-A patients. The gene encodes a 44 kDa protein which is a member of the WD-repeat family of proteins (Henning et al., 1995). WD-repeats are highly conserved repeating units that are frequently regulatory and occur in many proteins involved in processes such as cell division, gene transcription, transmembrane signalling and mRNA modification (for review see Neer et al., 1994). The *CSA* gene is able to complement the UV sensitive defect of cells from rodent complementation group 8 and hence is identical to *ERCC8* (Itoh et al., 1996). Recombinant *CSA* protein has been shown in one report to interact *in vitro* with *CSB* and the TFIIF subunit p44 (Henning et al., 1995). Another group however found *CSA* and *CSB* were associated with different high molecular weight complexes (van Gool et al., 1997).

1.3.3.2 CSB

The *ERCC6* gene was cloned by functional complementation of the UV sensitive hamster mutant UV61 from rodent complementation group 6 (Troelstra et al., 1990). This gene is able to correct the UV sensitivity and RNA synthesis defect of CS-B cells, and mutations were identified in a CS-B patient. This confirmed that *ERCC6* was the *CSB* gene. The gene is located on chromosome 10q11-q21 and encodes a 1493 amino acid protein with a predicted molecular weight of 168 kDa (Troelstra et al., 1992). Figure 1.4 is a diagrammatic representation of the *CSB* gene showing putative functional domains. Of particular note are the seven consecutive domains that are characteristic of members of the SWI/SNF family of putative helicases and ATPases.

Figure 1.5 shows the SNF2 superfamily of proteins with examples, particularly the SNF2-like family of NTP-binding proteins. Other members of the family are involved a wide variety of cellular processes, for example *MOT1* in transcriptional repression (Davis et al.,

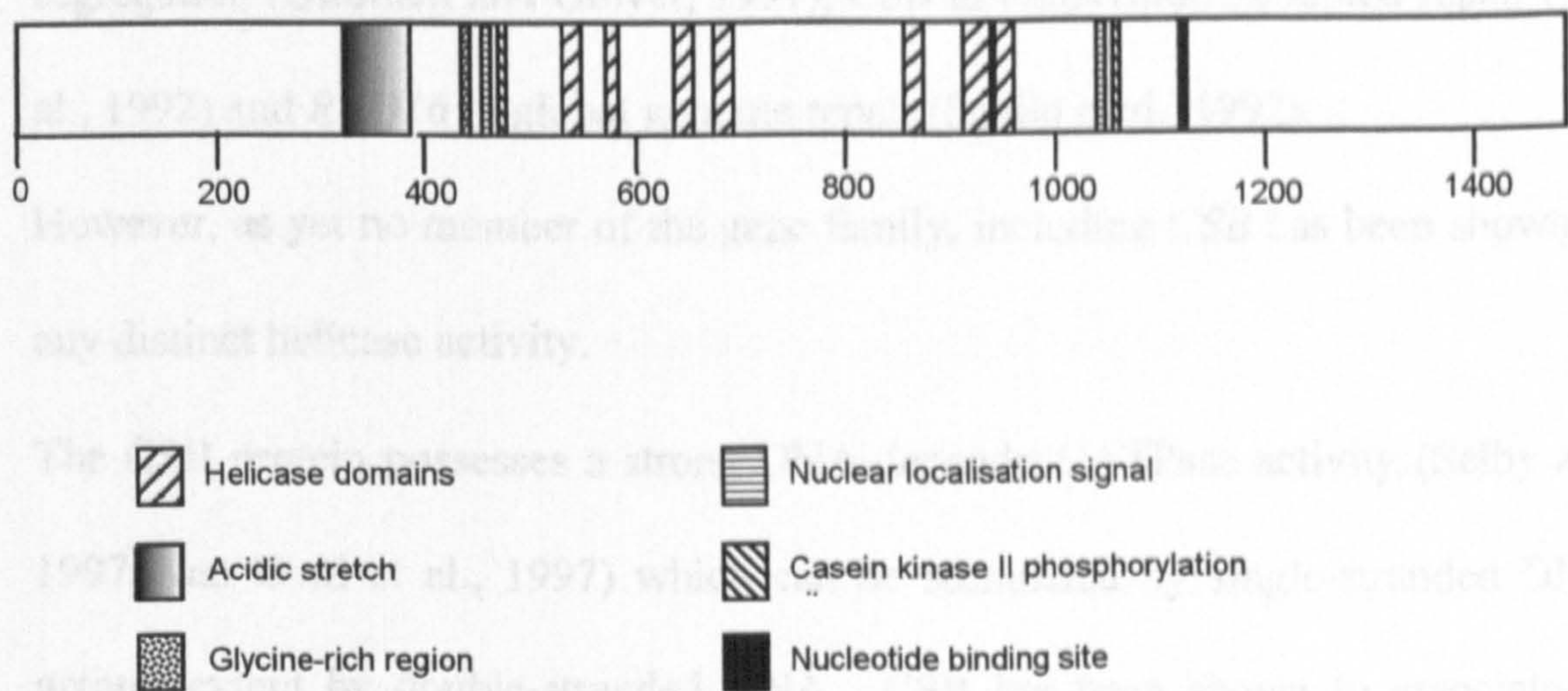


Figure 1.4 A diagrammatic representation of the CSB protein indicating putative functional domains.

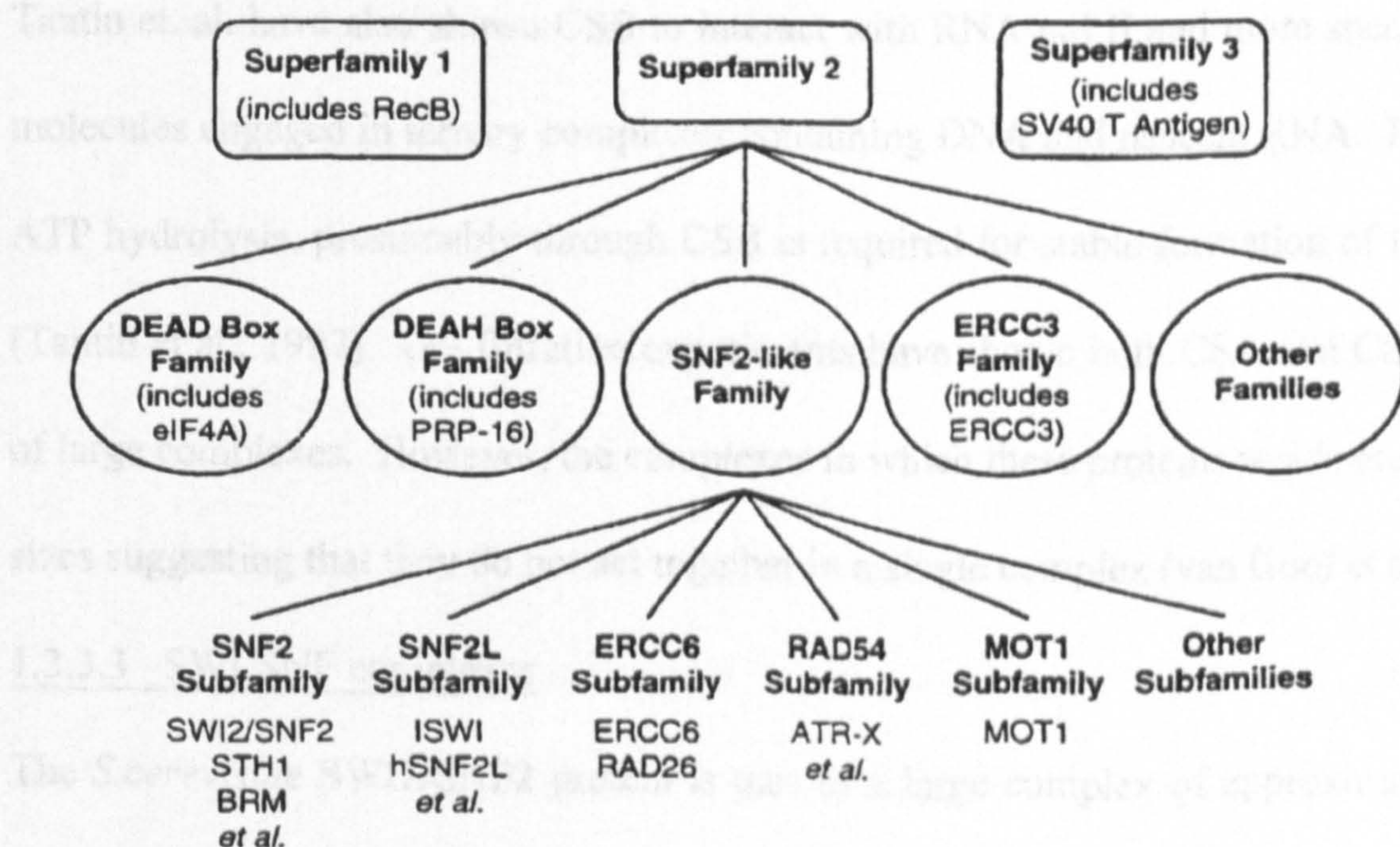


Figure 1.5 The SWI/SNF family of proteins (Pazin and Kadonaga, 1997)

1992), *RAD54* in recombination repair (Emery et al., 1991), *lodestar* in chromosome segregation (Girdham and Glover, 1991), *CSB* in transcription-coupled repair (Troelstra et al., 1992) and *RAD16* in global genome repair (Schild et al., 1992).

However, as yet no member of the gene family, including *CSB* has been shown to possess any distinct helicase activity.

The *CSB* protein possesses a strong DNA-dependent ATPase activity (Selby and Sancar, 1997; van Gool et al., 1997) which can be stimulated by single-stranded DNA or to a greater extent by double-stranded DNA. *CSB* has been shown to associate with *CSA* (Henning et al., 1995), *XPG* (Iyer et al., 1996), *XPA*, *XPB* and the p34 subunit of *TFIIIE* (Selby and Sancar, 1997) in experiments using either *in vitro* translated proteins, or GST pull-down assays. However, no such interactions were found when NER and transcription proficient cell-free extracts were used to determine possible interactions (van Gool et al., 1997). The only significant interaction found in these experiments was a proportion of approximately 10-15% of RNA pol II that bound tightly to *CSB* (van Gool et al., 1997). Tantin et. al. have also shown *CSB* to interact with RNA pol II and more specifically with molecules engaged in ternary complexes containing DNA and nascent RNA. Furthermore, ATP hydrolysis, presumably through *CSB* is required for stable formation of this complex (Tantin et al., 1997). Gel filtration experiments have shown both *CSA* and *CSB* to be part of large complexes. However, the complexes in which these proteins reside are of different sizes suggesting that they do not act together in a single complex (van Gool et al., 1997).

1.3.3.3 SWI/SNF complexes

The *S.cerevisiae* *SWI2/SNF2* protein is part of a large complex of approximately 2 Mda, comprising at least ten subunits including *SWI1*, *SWI3*, *SNF5* and *SNF6* (Cairns et al., 1994; Peterson et al., 1994). *SWI2/SNF2* is a positive regulator of transcription and shows *in vitro* DNA-dependent ATPase activity (Laurent et al., 1993). A link with chromatin structure was first noted due to mutations in core histones being able to suppress mutations

in the genes encoding the SWI/SNF complex (Hirchhorn et al., 1992; Kruger et al., 1995). Furthermore, *in vitro* DNaseI protection assays have shown that SWI/SNF can disrupt pre-assembled nucleosomes in an ATP-dependent manner (Owen-Hughes et al., 1996). The ATPase and nucleosome remodelling activity can be stimulated by single-stranded, double-stranded or nucleosomal DNA (Laurent et al., 1993).

A further nucleosome remodelling complex called RSC has been shown to exist in yeast. Many of the subunits are homologous to the subunits of SWI/SNF, including STH1 the ATPase homologue of SWI2/SNF2 that is consequently a member of the SNF2 gene family (Cairns et al., 1996). The RSC complex, in contrast to SWI/SNF contains subunits that are essential for mitotic growth, and as such it is likely to perform a function distinct from that of SWI/SNF.

Similar complexes also exist in humans, *Drosophila* and mouse with the SWI2/SNF2 homologues within these complexes being the most extensively studied. The *Drosophila* complex NURF is different from other chromatin remodelling complexes in the fact that its ATPase subunit ISWI is stimulated preferentially by nucleosomal DNA and not at all by free DNA. It has also been shown to be able to assist gene activation on chromatin templates at an early stage in transcription initiation (Mizuguchi et al., 1997). Therefore the *Drosophila* complex NURF performs a specific role in relaxing the chromatin structure at promoters allowing transcription to occur.

Human cells possess several SNF2 homologues with the genes hbrm and BRG1 being part of high molecular weight complexes that are able to disrupt nucleosomes *in vitro* (Kwon et al., 1994). The DNA repair protein CSB, as mentioned earlier is also a member of the SNF2 family and as such is a DNA-dependent ATPase (van Gool et al., 1997). The protein resides within a high molecular weight complex of at least 700 kDa. This may suggest a possible chromatin remodelling function, but as yet none of the other subunits have been

identified and the action of the complex has not yet been tested for its effect on nucleosomal structure.

1.3.3.4 CSB knockout mice

Knockout mice have been created for CSB, by mimicking the stop mutation at K337 found in the CS-B patient CS1AN (Troelstra et al., 1992). The mice generated with this mutation are viable and appear normal (van der Horst et al., 1997). Closer examination revealed the CS characteristics of UV sensitivity, inactive transcription-coupled repair, normal global genome repair and defective recovery of RNA synthesis after UV. However, the prominent clinical features of CS patients of growth failure and neurological dysfunction are exhibited only mildly in the mice. Interestingly, the mice also showed an increased susceptibility to skin cancer with multiple tumours appearing at 23 weeks in response to UV, compared with single tumours in wild-type mice at 27 weeks (van der Horst et al., 1997).

The theory presented for this significant difference between the human and mouse situation is that the TCR of CPDs helps to prevent carcinogenesis. Therefore the greater efficiency of GGR in humans is able to compensate for a defect in TCR better than in mice, and therefore humans do not present with skin cancer (van der Horst et al., 1997). Also, the life-span of humans is considerably longer than that of mice, although patients rarely live beyond their teens. Therefore, it is possible that if patients lived to a greater age they may demonstrate an increased incidence of skin cancer.

1.3.3.5 CSB and transcription

Cockayne syndrome cells fail to recover RNA synthesis after UV irradiation, but as described previously this does not always correlate with a defect in transcription-coupled repair (van Oosterwijk et al., 1996). Global genome repair in CS cells functions normally, therefore the majority of lesions will be repaired albeit more slowly. However, the phenotype in CS patients is very severe even in comparison to XP-A patients who are

defective in all NER. Therefore, many groups have predicted a transcription defect in these cells, since the repair defect cannot account for the severity of the clinical phenotype of the patients.

Extracts from Cockayne Syndrome cells from both complementation groups have been shown to have reduced levels of RNAPII transcription *in vitro* (Dianov et al., 1997). Balajee et. al. studied levels of transcription *in vivo* in intact and permeabilised CS-B cells. Intact CS-B fibroblasts and lymphoblasts showed a 50% reduction in uridine pulse labelling and therefore transcription when compared with both normal and XP-A cells. This reduction in transcription could be complemented by transfection of the *CSB* gene. Transcription in permeabilised cells was analysed in chromatin which was isolated under physiological conditions. The level recorded in CS-B cells was about 50% of that in normal chromatin. In this case transcription can be restored to normal levels by addition of normal cell extracts (Balajee et al., 1997). These data demonstrate a link between the *CSA* and *CSB* gene products and transcription, even in the absence of damage (discussed in greater detail in chapter 7).

Further evidence linking the stimulation of repair to transcriptional arrest was demonstrated by Wang et. al. with the use of triplex forming oligonucleotides. Specific oligonucleotides can be used to block transcription initiation and elongation through their ability to bind in the major groove of duplex DNA at polypurine-polypyrimidine sequences (Beal and Dervan, 1991). When mammalian cells are treated with triplex-forming oligos, mutations can be specifically induced in an SV40 vector contained within the cells. Such triplex-induced mutagenic activity cannot be detected in CS-B or XP-A cells suggesting a requirement for TCR in this process. Furthermore, the oligos that induced mutagenic activity in normal cells also inhibited transcription and stimulated repair (Wang et al., 1996). Again a process or assay that results in transcriptional arrest requires active repair and the CSB protein.

The CSB protein has also been shown to enhance RNAPII elongation of an artificial undamaged template by a factor of three (Selby and Sancar, 1997). A template containing a T-T cyclobutane dimer results in a strong block to transcription, and addition of CSB to this template induces the addition of a single nucleotide to the nascent transcript. This suggests that CSB is able to recognise either the damage or the stalled polymerase and either induces the addition of the nucleotide as a signal for repair, or for the restart of transcription.

Selby and Sancar have also looked at the effects of CSB and TFIIS on a blocked RNA polymerase molecule. TFIIS can stimulate the backward movement of RNAPII and transcriptional read-through, possibly by allowing access of the repair machinery to a lesion site (Aso et al., 1995; Reines et al., 1996). This backward movement also involves the digestion of the nascent transcript by a blocked polymerase. This has led to the theory that TFIIS may be involved in TCR by acting together with CSA and CSB in signalling damage and recruiting repair machinery, through association with a stalled polymerase. Addition of CSB to a stalled polymerase can counteract the TFIIS-induced transcript shortening (Selby and Sancar, 1997). Therefore, it is unlikely that TFIIS works with CSB as a TRCF, since they appear to work at odds with respect to the nascent transcript. Also, disruption of yeast TFIIS in *S.cerevisiae* had no overt effect on TCR (Verhage et al., 1997), demonstrating that any involvement this molecule may have in TCR is not essential for the pathway, at least in yeast.

Mu and Sancar have developed an *in vitro* model system using purified human excision repair factors and defined substrates. They have attempted to mimic NER with a DNA duplex substrate carrying a T-T dimer, and TCR with a transcription-like bubble that terminates at a cyclobutane thymine dimer. The dimer is removed three times faster from the bubble substrate than from duplex DNA. The faster repair from the bubble substrate is independent of the XPC protein (Mu and Sancar, 1997) which is now thought to be

required for the recognition of damage during NER. Therefore the lack of requirement for XPC in TCR suggests that damage recognition in this reaction occurs by a separate mechanism, possibly through a stalled polymerase molecule or CSB.

1.3.3.6 RNA polymerase II ubiquitination

Recent experiments by Bregman and co-workers have suggested a link between the CS genes and ubiquitination of RNA polymerase II. After UV or cisplatin treatment, a fraction of the large subunit of RNAPII was ubiquitinated, but this process did not occur in CS-A or CS-B cells. Ubiquitination could be restored by introduction of the CS cDNA constructs (Bregman et al., 1996).

In normal cells UV is able to induce a transient reduction in the levels of the large subunit of RNAPII, presumably as a consequence of ubiquitination. Normal levels of RNAPII return after 16-24 hours through new protein synthesis (Ratner et al., 1998). In contrast to the lack of ubiquitination and degradation of RNA pol II in CS cells, repair deficient XP cells show a sustained reduction in RNAPII large subunit, indicating an inability to induce new protein synthesis. UV-induced reduction of RNAPII large subunit can be inhibited in both normal and repair-deficient cells by proteasome inhibitors, demonstrating that the ubiquitination is a signal for degradation (Ratner et al., 1998).

Ubiquitination also appears to be linked to the phosphorylation state of the RNAP C-terminal domain (CTD), which as described earlier is highly phosphorylated in elongating RNAPII complexes. In all cells, a ubiquitinated RNAPII large subunit has a hyperphosphorylated CTD, but the portion of RNAPII whose levels were most quickly reduced were relatively hypophosphorylated. However, the ubiquitinated residues do not map to the CTD (Ratner et al., 1998), showing that phosphorylation is not a signal for ubiquitination and vice-versa. Although ubiquitination is not triggered by phosphorylation, the phosphorylation state of RNAPII-CTD is a good indicator of the transcriptional status of the polymerase.

It has been suggested that stalled RNAPII complexes backtrack to allow repair of the blocking lesion, then the RNA synthesis restarts. The above data contradicts this theory, suggesting that RNAPII stalled at a lesion is targeted for degradation through ubiquitination. Therefore rather than transcription continuing after repair, it is likely that the transcript is aborted. This is in agreement with the reaction in *E.coli*, whereby a stalled RNAP is displaced by TRCF (Pavco and Steege, 1990). Lack of ubiquitination in CS cells is consistent with the idea that CSA and CSB act in a similar fashion as TRCF, removing stalled RNA polymerase molecules. Therefore CSA and CSB might be involved in the targeting of ubiquitin conjugating enzymes to stalled RNAPII molecules, leading to their subsequent removal and degradation.

1.3.3.7 Transcription-coupled repair – a link with mismatch repair

The mismatch repair pathway is responsible for the correction of single base mismatches and small insertions or deletions that can occur during replication (for review see Friedberg et al., 1995). Patients defective in mismatch repair are associated with the hereditary form of colon cancer HNPCC (hereditary non-polyposis colon carcinoma).

In *E.coli* the major proteins involved in this process are MutH, MutS and MutL, and recently it has been suggested that these proteins are also involved in transcription-coupled repair. This theory arose from the observation that residual MFD seen in transcription-inhibited *mfd-1* cells was of a very similar level to that occurring in the mismatch repair mutants *mutS* and *mutL* (Li and Bockrath, 1995). Also, levels of TCR in the *lacI* gene were very similar in both *mfd-1* and mismatch repair deficient cells (Mellon and Champe, 1996). However, Selby and Sancar could not detect a TCR defect in *mutL* or *mutS* cells and were able to reconstitute TCR *in vitro* without the addition of these proteins (Selby and Sancar, 1993).

Reduced transcription-coupled repair is also seen in mismatch deficient human cells such as tumour cell lines and HNPCC-derived lymphoblastoid cell lines (Mellon et al., 1996).

More specifically, defects in the *MutS* and *MutL* homologues *hMSH2* or *hMLH1* result in reduced TCR of UV damage without affecting the levels of overall genome repair (Leadon and Avrutskaya, 1997). The same however, is not true in the yeast *S.cerevisiae*, with normal TCR observed in cells defective for *MSH2*, *MLH1*, *PMS1* or *MSH3*. Additionally, no effect can be seen through mutation of *MSH2* in either a global genome ($\Delta RAD7$) or transcription-coupled ($\Delta RAD26$) repair deficient background (Sweder et al., 1996).

The only circumstances in yeast where mismatch repair deficiency effects transcription-coupled repair is in the removal of oxidative damage. *MSH2* deficient yeast are unable to remove thymine glycols from the transcribed strand of the *RPB2* gene (Leadon and Avrutskaya, 1998). The same defect is observed in double mutants of the *MutL* homologues *MLH1* and *PMS1*, but not with mutation of either alone (Leadon and Avrutskaya, 1998). Transcription-coupled repair of thymine glycols in human cells does not have the same requirement as TCR of UV damage, with respect to the involvement of mismatch repair proteins. Cells defective for *hMSH2* are unable to remove oxidative damage from the transcribed strand of active genes, whereas repair of such regions is unaffected by mutation of *hMLH1* (Mellon et al., 1996).

The above data suggest that the involvement of mismatch repair in transcription-coupled repair is not evolutionarily conserved.

CHAPTER 2 MATERIALS AND METHODS

2.1 Cell Culture Techniques

2.1.1 Primary human fibroblasts

Primary human fibroblasts were cultured in Eagle's MEM (Gibco BRL) supplemented with 12.8 ml 7.5% sodium bicarbonate (Gibco BRL), 4.5 ml each of 200mM L-glutamine (Gibco BRL) and 10000 U/ml penicillin/10000 µg/ml streptomycin (Gibco BRL), and 15% foetal calf serum. Cultures were grown under 5% CO₂.

2.1.2 CHO cells

CHO cells were cultured in Eagle's MEM supplemented with 12.8 ml sodium bicarbonate, 4.5 ml each L-glutamine, penicillin/streptomycin and 100X non-essential amino acids (Gibco BRL) and 10% foetal calf serum. Cultures were grown under 5% CO₂.

2.1.3 Sub-culturing and harvesting

Cells were trypsinised and spun down through 5 ml of medium. The pellet was resuspended in an appropriate volume and counted on a haemocytometer. Large flasks for growth of fibroblasts were re-seeded with a minimum of 10⁶ cells. CHO cells were grown up from the cells remaining after trypsinisation.

2.1.4 Transfection

UV61 CHOs were plated on 10 cm dishes at a density of 5x10⁵ cells per dish and were grown for two days. On the second day, 3 ml of fresh medium was added to the cells. Forty microlitres of 10 mg/ml polybrene was added and washed over the cells followed by 5 µg of the DNA to be transfected. The dishes were incubated at 37°C for 6 hr. Following

incubation, the cells were treated with 20% DMSO in serum-free medium for 4 min. The cells were then washed three times in serum-free medium and returned to standard culture medium overnight. The medium was changed the following day. Two days post-transfection the cells were trypsinised and each culture seeded back into three new 10 cm dishes and selection was applied using 1 mg/ml G418. The cells were cultured, refreshing the selection medium every 4-5 days until colonies formed, typically after around 15 days. Colonies were isolated using cloning rings, adding trypsin to the colony within the ring. The detached colony was transferred to a bijoux containing 1 ml of medium and split between a small flask and a single well in a six-well plate.

2.1.5 Freezing

Trypsinised cells were suspended in 10% DMSO in standard culture medium and frozen at a density of 2×10^6 / ml.

2.1.6 UV survival-Method I

Cells were plated at densities of 10^2 , 2×10^2 , 4×10^2 and 8×10^2 per 6 cm dish corresponding to increasing doses. Four dishes were prepared at each density and left overnight to attach. The dishes were irradiated with UV-C at 0, 3, 6 and 9 J/m² at a rate of 0.5 J/m²/sec and the medium replaced. After 7 days of colony growth, the cells were fixed, stained and the number of colonies counted.

2.1.7 UV survival-Method II

Cells were plated at 10^5 per 6 cm dish, plating one dish for each dose and were left overnight to attach. The dishes were irradiated as for method I with 0, 3, 6 and 9 J/m², then trypsinised and counted. The cells were plated at the densities indicated in method I, preparing four dishes for each dose. After 7 days of growth, colonies were counted and

survival expressed as a percentage of un-irradiated colony numbers after correction for plating efficiency.

2.1.8 RNA synthesis after UV-C irradiation

Ten 3 cm dishes were prepared for each cell line, plating 10^5 cells per dish. Each experiment contained control cell lines of wild-type and mutant for transcription-coupled repair as well as the cells of interest. The control cells in this case were the CSB hamster mutant line UV61 and the parental line AA8. After overnight incubation, the dishes were irradiated at doses of 0, 3, 6, 9 and 12 J/m² with duplicate dishes for each dose. The cells were allowed to recover for 7 hr and then pulse labeled for 30 min in 2 ml of medium containing 5 μ Ci/ml specific activity ³H-uridine. The medium was removed, the cells washed with PBS and lysed with 0.5 ml of 2% SDS. The cells were scraped off the dishes with silicon rubber and 100 μ l duplicate samples per dish were each spotted onto 3 cm 3MM filter strips. The filters were treated in 5% trichloroacetic acid for 5-10 min, then washed twice in IMS and allowed to air dry. Incorporation of ³H was measured by scintillation counting and post-irradiation RNA synthesis expressed as a percentage of levels in un-irradiated cells.

2.2 Molecular Methods

2.2.1 Total RNA extraction

Human skin fibroblasts were harvested from a confluent large flask containing between 5×10^6 and 10^7 cells. The monolayer was trypsinised and spun down through 5 ml medium. The pellet was resuspended in a small volume of PBS and transferred to an eppendorf. 500 μ l of Solution D (4M guanidinium thiocyanate, 25mM sodium citrate, 0.5% sarcosyl; 0.1M 2-mercaptoethanol), 50 μ l of 2M sodium acetate and 500 μ l of phenol:chloroform

(50:50) were added sequentially, mixing after each addition. The solution was spun at 15,000 rpm for 5 min then the upper aqueous phase transferred to a fresh tube. The RNA was precipitated with 500 µl of isopropanol, spinning for 10 min at 15,000 rpm. The pellet was washed with 70% ethanol and resuspended in 50 µl filter sterile double distilled water. The RNA was used directly for cDNA synthesis or stored in ethanol at -80°C.

2.2.2 cDNA synthesis

cDNA synthesis was carried out in three sections using the antisense PCR primers for fragments II (EC176 25-mer), IV (E6-2) and VI (EC179 no.6) (see Table 2.1 and Appendix I). Primers were at a concentration of 25 ng/µl and 2 µl of this solution was mixed with 8 µl of RNA. The reaction was heated to 80°C for 10 min, then put on ice briefly and mixed with 4 µl first strand synthesis buffer (250mM Tris-HCl pH8.3, 375mM KCl, 15mM MgCl₂), 2 µl 10mM dNTPs (Gibco BRL), 2 µl 0.1M DTT and 1 µl MMuLV Reverse Transcriptase (10U/µl, Pharmacia). The reaction was incubated at 37°C for 2 hr then made up to a volume of 40 µl ready for PCR.

2.2.3 PCR

The gene was amplified in six pieces, using the cDNAs, primers and cycling conditions outlined in Table 2.1. Each reaction was carried out in a 50µl volume comprising of 5 µl Px10 buffer (100mM Tris-HCl pH9, 500mM KCl, 15mM MgCl₂, 0.1% gelatin), 2.5 µl 2.5mM dNTP mix, 1µl of each primer at 20 pmol/µl, 1 µl of 0.5 U/µl Red Hot Polymerase (Advanced Biotechnologies), 5 µl cDNA and SDW. Reactions were overlaid with 50 µl mineral oil before amplification. PCR products were visualised on a 0.8% agarose gel.

Fragment	cDNA primer	5' primer	3' primer	Cycling conditions
I (1134 bp)	II	F1 (EC179 no.1)	R1 (EC179 no.2)	94°C 5min x 1 cycle 94°C 1min 65°C 1.5min x 40 cycles 72°C 3min
II (889 bp)	II	F2 (EC176 25-mer)	R2 (EC176 25-mer)	94°C 5min x 1 cycle 94°C 1min 63°C 1.5min x 40 cycles 72°C 3min
III (671 bp)	IV	F3 (E6-5)	R3 (E6-6)	94°C 5min x 1 cycle 94°C 1min 67°C 1.5min x 40 cycles 72°C 3min
IV (1146 bp)	IV	F4 (E6-1)	R4 (E6-2)	94°C 5min x 1 cycle 94°C 1min 67°C 1.5min x 40 cycles 72°C 3min
V (1482 bp, 960 bp)	VI	F5 (E6-3, E6-9)	R5 (E6-4, E6-10)	94°C 5min x 1 cycle 94°C 1min 65°C 1.5min x 40 cycles 72°C 3min
VI (928 bp)	VI	F6 (EC179 no.5)	R6 (EC179 no.6)	94°C 5min x 1 cycle 94°C 1min 60°C 1.5min x 40 cycles 72°C 3min

Table 2.1 The cDNAs, PCR primers and amplification conditions used for the PCR amplification of CSB cDNA in six fragments.

2.2.4 Re-amplification

The first round PCR product was run on a 2% low melting point agarose gel and the desired band excised from the gel. The agarose was diluted with 200-500 µl TE depending on the amount of DNA and melted at 65°C for use as a template in further rounds of amplification. Second round amplifications were carried out as before, replacing the sense primer with a nested biotin-tagged primer. The oligonucleotides were synthesised with an amino-linker on the 5' end and removed from the column with 10% ammonia at 55°C overnight. The solution was run through a NAP10 column and the oligonucleotide eluted

in 1 ml TE. For addition of the biotin, half of the oligonucleotide stock was precipitated and resuspended in 0.1M HEPES. This was mixed with biotinyl succinimide overnight at room temp and purified on a NAP10 column.

2.2.5 Selection of oligonucleotides

Oligonucleotides were obtained from Rotterdam for amplification of the full length open reading frame of the *CSB* gene, giving amplification in six fragments. Some of the primer pairs gave good or adequate products, others however gave no amplification irrespective of the conditions employed during cycling (see chapter 3). Replacement primers were designed to cover the regions that did not amplify with the original sets. The primers were designed as 24-mers on the Oligo 5.0 computer program which is able to screen chosen sequences for mis-priming, primer dimers and hairpins. It is also possible to predict melting temperatures and annealing temperatures, although cycling conditions need to be checked fully as is discussed in the section below. In total, new oligonucleotides were designed for fragments III, IV and V.

2.2.6 Strand separation

The second round PCR products were diluted with TE and mixed with 50 µl streptavidin-coated magnetic beads for 1-2 hr. The beads had been washed twice with 2x Bind and Wash buffer (10mM Tris-HCl pH7.5, 1mM EDTA, 2M NaCl). The beads were then washed twice with Triton Wash Solution (TWS-0.17% Triton X-100, 500mM NaCl, 10mM Tris-HCl pH7.5, 1mM EDTA pH8) and resuspended in 32 µl TE. The DNA was denatured by addition of 8 µl of 2M NaOH/4mM EDTA with 5 min incubation at room temp. At this stage the biotinylated sense strand remains attached to the magnetic beads and the other strand is in the supernatant. The supernatant was run on an agarose gel to assess the amount of ssDNA in the supernatant, giving an estimate of the ssDNA still on

the beads and therefore the binding efficiency. The beads were again washed twice with TWS and once with TE, followed by resuspension in SDW ready for sequencing.

2.2.7 Sequencing

Dynabead bound ssDNA and ssDNA plasmid preps were subjected to sequencing carried out according to the Sequenase version 2.0 kit (Amersham). Reactions were run on denaturing polyacrylamide gels for approximately two hours at a constant power of 60W, dried and exposed to autoradiographic film overnight.

2.2.8 DNA purification

DNA was purified according to the Glassmax DNA Isolation Matrix kit (Gibco BRL) following gel purification.

2.2.9 Standard techniques

Ligations, minipreps, ssDNA preparation, restriction digests and transformations were carried out by standard procedures as described by Sambrook et. al. 1989.

2.2.10 Maxipreps

Two procedures were used for the large scale preparation of plasmid DNA. A 200 ml overnight culture was harvested by spinning at 5,000 rpm for 5 min. The pellet was resuspended in 10 ml Solution I (50mM glucose, 25mM Tris-HCl pH8, 10mM EDTA), followed by sequential addition of 20 ml Solution II (0.2M NaOH, 1% SDS) and 10 ml Solution III (60% 5M KCl, 11.5% acetic acid). The solution was incubated on ice for 30 min then spun down at 10,000 rpm for 10 min. The supernatant was filtered through gauze into 100 ml of ethanol and the DNA spun down (10,000 rpm, 10 min).

2.2.10.1 Method I

The DNA was resuspended in 5 ml 5xTE and mixed with 6 g Caesium chloride and 500 µl of 1 mg/ml ethidium bromide. This was spun at 48,000 rpm for at least 17 hr and the plasmid band removed. The ethidium bromide was removed by iso-amyl alcohol extraction and the DNA ethanol precipitated. The pellet was resuspended in 500 µl TE and the concentration determined using the Pharmacia Genequant.

2.2.10.2 Method II

The DNA was resuspended in 400 µl of TE and treated with 4 µl of 1 mg/ml boiled RNase for 15 min at 37°C. The solution was precipitated with 200 µl of 2.5M NaCl/20% polyethylene glycol on ice for 15 min and spun down at 15,000 rpm for 10 min. The pellet was resuspended in 400 µl of TE and extracted with phenol twice followed by two chloroform/iso-amyl alcohol (24:1) extractions. The DNA was ethanol precipitated and resuspended in 1 ml TE and the concentration determined as for method I.

2.2.11 Site-directed mutagenesis I

The plasmids pcBlSE6 (WT CSB within pBluescript, from C.Troelstra) and pCI-WT (WT CSB within pCI-neo, Promega) were transformed into the ung-, dut- strain CJ236 to produce uracil containing ssDNA. The mutagenesis primers were designed to carry the change in the middle of a 26-mer in the case of single base changes. For greater differences or deletions, oligonucleotides of up to 40 nt were synthesised.

Mutagenesis primers were phosphorylated by mixing 2 µl of 10 ng/µl primer with 20 µl 5x Ligation buffer with ATP (Gibco BRL) and 5U Polynucleotide kinase (Pharmacia) in a total volume of 100 µl. The mixture was incubated at 37°C for one hour. For the primer annealing reaction, 10 µl of phosphorylated primer was mixed with 100 ng ssDNA and 2 µl 10 x annealing buffer (200mM Tris-HCl pH7.5, 20mM MgCl₂, 500mM NaCl) and heated to 65°C for 3 min followed by 10 min at 37°C. For synthesis and ligation of the

complementary strand, 2 µl 10 x synthesis buffer (175mM Tris-HCl pH7.5, 37.5mM MgCl₂, 21.5mM DTT, 8mM ATP, 4mM dNTP), 1 µl BRL ligase (1 U/µl) and 1 µl T7 DNA polymerase were added and the reaction incubated at 37°C for 2 hr. Half of the reaction mix was transformed into competent DH5αF', where uracil containing DNA will be degraded and only the newly synthesised mutant DNA will result in ampicillin resistant colonies. Individual colonies were picked and checked for presence of the mutation either by restriction analysis or sequencing. All constructs were sequenced through the entire CSB open reading frame to check for polymerase errors.

2.2.12 Site-directed mutagenesis II

Site-directed mutagenesis reactions were also carried out using the Stratagene Quikchange mutagenesis kit. Primers carrying the desired mutation were designed complementary to both strands of the DNA. In general 25 µg of double stranded pCI-WT plasmid was mixed with 125 ng of each primer, 2 µl 10 x buffer, 1.5mM dNTPs and 2 units of Pfu polymerase. The mixture was overlayed with 25 µl of mineral oil and a PCR like reaction was carried out annealing at 55°C for 1 min and extension for 20 min at 65°C (2 min per kb of plasmid). After amplification, 1 unit of *DpnI* restriction enzyme was added and the reactions incubated at 37°C for 1 hour. Between 1 and 10 µl of the reaction was transformed into the supercompetent *E.coli* cells XL-1 blue (Stratagene).

2.3 Biochemical Techniques

2.3.1 Whole cell extracts I

CHO cell lines were grown to confluency on 10 cm dishes, trypsinised and washed three times with ice cold PBS. The cell pellet was finally resuspended in 300 µl ice cold PBS,

vortexed for three short bursts of 10 sec, then sonicated twice for 30 sec, placing onto ice between each burst.

2.3.2 Whole cell extracts II

Cell pellets were washed three times in PBS then resuspended in 200 µl of buffer (50mM NaF, 20mM HEPES pH7.8, 450mM NaCl, 25% glycerol, 0.2mM EDTA, 0.5mM DTT and protease inhibitor cocktail, Boehringer). The cell suspension was then subjected to snap freezing on dry ice/ethanol, followed by rapid thawing at 30°C. The cycle of freeze thawing was repeated twice more. Cell debris was spun down for 5 min at 12,000 rpm at 4°C. The supernatant was aliquoted into 10 µl samples and snap frozen immediately on dry ice/ethanol.

2.3.3 Cell fractionation I

Cell fractions were prepared as described by Krude et. al. 1997. Cells were grown to near confluence on 10 cm dishes and washed once with PBS. They were then washed with ice cold hypotonic buffer (20mM potassium-HEPES, 5mM potassium acetate, 0.5mM MgCl₂, 0.5mM DTT), then replaced with 10 ml of the same buffer and allowed to swell for 10 min at 4°C. After swelling all excess buffer was removed and the cells scraped off into the residual liquid. The suspension was transferred into a dounce homogeniser and the cells disrupted with 25-50 strokes. The suspension was transferred to an eppendorf and the nuclei spun down at 4,000 rpm for 5 min at 4°C. The supernatant containing the cytoplasmic extract was transferred to a fresh tube. The pelleted nuclei were washed three times with PBS and finally resuspended in 100 µl of hypotonic buffer with 0.4M NaCl. The nuclei were incubated on ice for 90 min and nuclear material pelleted at 13,000 rpm for 30 min at 4°C. The supernatant contains the nuclear extract and the pellet nuclear debris.

2.3.4 Cell fractionation II

Cell fractions were prepared as described by Rathmell and Chu, 1994. Cell pellets were washed several times in PBS, once in wash buffer (10mM HEPES pH7.5, 10mM KCl, 0.5mM DTT), and finally in 40 µl wash buffer with 4 µl 10% NP-40. Cells were lysed in this buffer at 4°C with mixing for 10-20 min, checking for lysis periodically. Nuclei were pelleted at 13,000 rpm for 2 min at 4°C and the resulting supernatant is the cytoplasmic fraction. The nuclei were then resuspended in 40 µl high salt buffer (20mM HEPES pH7.5, 500mM NaCl, 0.2mM EDTA, 0.5mM PMSF, 0.5mM DTT, 1.5mM MgCl₂, 20% glycerol) and extracted for 15 min on ice. Nuclear debris was pelleted at 13,000 rpm for 10 min at 4°C, with the supernatant being the nuclear extract.

2.3.5 SDS PAGE

Proteins were run on 8% denaturing polyacrylamide gels with 0.1% SDS in 1x running buffer (25mM Tris, 0.19M glycine, 0.1% SDS). Gels were run using the Bio-Rad mini protean system until the 45 kDa marker reached the bottom of the gel.

2.3.6 Western Blotting

SDS-Page gels were transferred to a tank containing cold blotting buffer (3 g/l Tris, 145 g/l glycine, 0.01% SDS) and soaked for 5-10 min. The proteins were transferred onto Nitrobind nitrocellulose transfer membrane (MSI) using the Bio-Rad MiniProtean wet blotting system, either overnight at 15 V or for 3 hr at 100 V. The blot was then blocked for at least 1 hr in PBS with 10% milk, 0.1% Tween.

2.3.6.1 HA antibody

Blocked blots were mixed with either the commercially obtained HA (BabCo) or HA purified from 12CA5 hybridoma cells at 1:1000 dilution in PBS with 2% milk, 0.1%

Tween at room temperature for 1-2 hr. Blots were then washed three times for 10 min in the PBS-milk-Tween solution, then incubated with horseradish peroxidase-conjugated mouse immunoglobulins (Dako) at 1:2500 dilution in PBS-milk-Tween for 1 hr at room temperature. Blots were then washed a further 3 x 10 min and finally with PBS-Tween to remove the milk. Proteins were visualised using the ECL detection system and autoradiography.

2.3.6.2 CSB antibody

Blocked blots were washed with PBS with 0.1% Tween to remove excess milk and transferred to a 15 ml tube containing 4 ml of polyclonal CSB antibody (van Gool et al., 1997) diluted 1:400 in PBS-Tween. The blot was incubated with the antibody at 4°C overnight with rolling. Blots were washed 3 x 10 min in PBS-Tween then incubated for 1½ hr at 4°C with 1:1000 dilution (PBS-Tween) of alkaline phosphatase conjugated rabbit immunoglobulins (BioSource International). After incubation with the secondary antibody, blots were again washed 3 x 10 min in PBS-Tween, finally washing in AP buffer (100mM Tris pH9.5, 100mM NaCl, 5mM MgCl₂) for 5 min. Proteins were visualised by adding 10 ml of AP buffer containing 111 µl nitroblue tetrazolium chloride (one tablet dissolved in 1 ml water, Sigma) and 11 µl bromo-chloro-indolyl-phosphate (one tablet dissolved in 500 µl DMF, Sigma). The reaction was stopped by removing the buffer and washing the filter with either water or TE.

2.3.7 Purification of HA antibody

Supernatant from 12CA5 hybridoma cells was harvested and spun down to remove cell debris. Saturated ammonium sulphate solution was added to the supernatant to give a final concentration of around 50%. The precipitate was spun down at 10,000 rpm for 5 min and resuspended in ice cold PBS. The suspension was dialysed overnight into PBS with several changes of buffer. The solution was run through a 1 ml HiTrap protein G column

(Pharmacia) and the antibody eluted with 1M glycine (pH 2.5). Immediately after elution the antibody fractions were neutralised with 1M Tris. Fractions were tested for protein content using the Bio-Rad protein assay and checked on a Coomassie stained SDS-PAGE gel for immunoglobulin bands. The peak fractions were pooled and dialysed again into PBS.

3.1 Introduction

There are a number of different techniques available for the sequencing of cloned genes to identify mutations occurring in patients suffering from genetic syndromes. Analysis of such genes involves extracting RNA from cultured primary fibroblasts, performing RT-PCR and direct sequencing. At the start of this study, we were provided with a series of primers and conditions by C. Troelstra, Rotterdam, that she had used to amplify the *CSB* gene from human cells in six fragments of around 1kb. However, different thermal cyclers produce different conditions and variation in starting materials can result in sub-optimal conditions. In my hands, the conditions used by Troelstra were not suitable for the amplification of *CSB*. Some of the primer pairs gave good or adequate products, others however gave no amplification irrespective of the conditions employed during cycling. Replacement primers were designed to cover the regions that did not amplify with the original sets. The primers were designed as 24-mers on the Oligo 5.0 computer program that is able to screen chosen sequences for mis-priming, primer dimers and hairpins. It is also possible to predict melting temperatures and annealing temperatures, although cycling conditions need to be checked fully as discussed in the section below. In total, new oligos were designed for fragments III, IV and V. There were also several methods being employed in the laboratory for RNA extraction, cDNA synthesis and direct sequencing. To determine the most efficient way to screen Cockayne syndrome type B patients for mutations it was necessary to check all methods for suitability and optimise those selected. Ultimately I was able to amplify *CSB* using a similar strategy to that used by Troelstra et.al., 1992 through the six fragments and primers shown in figure 3.1.

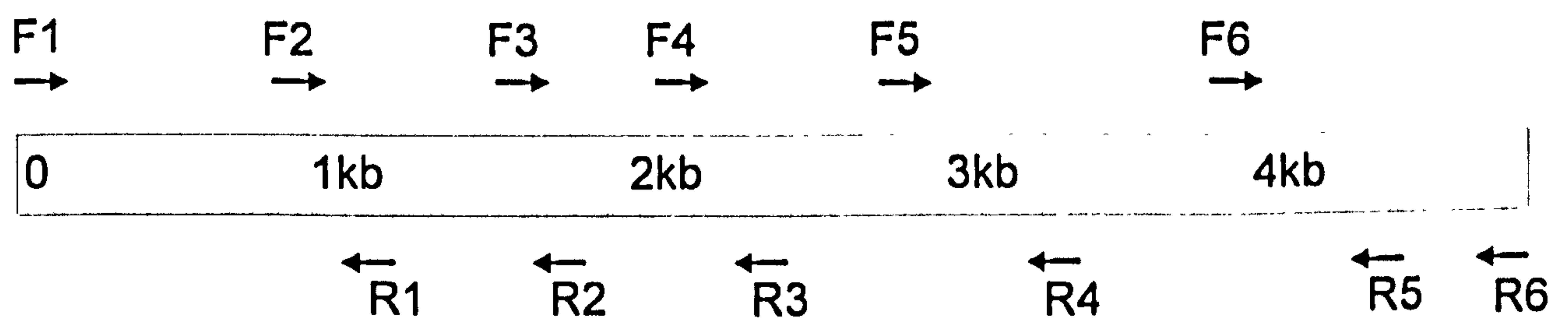


Figure 3.1 The location of the PCR primers finally used to amplify the CSB cDNA in six fragments.

3.2 RNA extraction

RNA extraction can be performed with two aims, to produce total RNA or poly(A) mRNA with the latter a more complex and time consuming procedure. The cDNA synthesis and PCR amplification methods should be sensitive enough for total RNA to be sufficient. Total RNA extraction has been carried out regularly in the laboratory using two procedures, cell lysis with non-ionic detergent and guanidinium thiocyanate extraction.

3.2.1 Non-ionic detergent lysis

This method involves lysing the cell pellet using the non-ionic detergent NP-40 and then spinning out the nuclei. The cytoplasmic supernatant contains the RNA, whereas the genomic DNA remains in the nuclear pellet. This method therefore allows both RNA and DNA preps to be performed on the same cell pellet.

3.2.2 Guanidinium Thiocyanate

This method uses the chemical guanidinium thiocyanate to break open both the cells and nuclei. It is also able to protect the RNA from degradation during the lysis procedure through inhibition of RNases. The released RNA is then purified by phenol extraction and precipitated which also removes guanidinium from the preparation.

3.2.3 Results and discussion

Both methods of extraction produced around around 50 µg total RNA from approximately 5×10^6 primary fibroblasts. When run on either a formaldehyde or sterile standard agarose gel, the RNA was not degraded and the 28S and 18S ribosomal RNA bands were clearly visible at the 2:1 ratio expected (Figure 3.2). It was also of sufficiently good quality to give cDNA that amplified reproducibly in a PCR reaction. Both methods of extraction were used during this project with equal success.

3.3 cDNA synthesis

The cDNA synthesis reaction, a standard procedure was carried out as outlined in chapter 2. The variable part of the procedure was the type of priming method chosen, resulting in very different amplification targets. There are three main types of priming used for cDNA; random hexamers, oligo dT and gene specific primers. To test the efficiency of cDNA synthesis it was necessary to use PCR of a control *hprt* DNA fragment and of the desired gene. It is important to use the target gene to test the cDNA, as genes all have different compositions. For example, GC rich regions and palindromic sequences will result in secondary structure and therefore cause problems in elongation.

3.3.1 Oligo dT

Oligo dT primers anneal to the poly(A) tail of mRNA and extend upstream from the single priming site. This technique works very efficiently for the short *hprt* gene and the 3' end of *CSB*. However, the efficiency of synthesis decreases with length of extension either due to the RNA not being intact or the cDNA synthesis not reaching the full length of the RNA transcript. This makes amplification of the 5' regions particularly difficult with large genes such as the 4.7 kb *CSB* cDNA.

3.3.2 Random hexamers

Random hexamers are 6 nt primers made up of random nucleotides which, because of their random nature are able to anneal at many different positions on the RNA. This method of priming is useful for long genes giving priming throughout the length of RNA, so that elongation along the full length of the transcript is not required. The main drawback with this method is that the primers consist of a wide variety of sequences and are able to anneal in a large number of places within the RNA. Therefore, the cDNA made from the target sequence will only make up a small proportion of the total cDNA synthesised.

3.3.3 Specific priming

This method involves the use of antisense primers specific for the gene of interest to prime the synthesis of cDNA. The number of primers needed depends on the size of the transcript and the efficiency of the elongation reaction. Generally, to be certain of high yield quality cDNA each reaction should not be expected to extend further than 2-3 kilobases.

3.3.4 Results and discussion

The cDNAs were initially tested for amplification of a control fragment from the *hprt* gene and the fragment VI at the 3' end of *CSB*. Any cDNAs that gave products in each of the above test PCRs was checked for amplification of the rest of the *CSB* fragments.

Random hexamer primed cDNAs failed to give products with either of the test fragments (Figure 3.3) and this procedure was not pursued.

Oligo dT was the second primer tested for cDNA synthesis. All cDNAs primed this way amplified well with the control fragment and *CSB* fragment VI (Figure 3.3). However, the farther upstream in the *CSB* gene the PCR primer pairs are situated, the weaker the amplification and the more variable the product (Figure 3.4). The oligo dT primed cDNAs

tested also failed to give PCR products for fragments I and II at the 5' end of the gene. Figure 3.4 shows the bands produced for fragment II. The main drawback to using oligo dT primed cDNA for amplification of *CSB* result from the gene being 4.7 kb. With a large gene the size of *CSB*, the cDNA needs to extend a long way from a single primer. Therefore, the RNA needs to be of high quality and the cDNA synthesis procedure has to be very efficient. Sub-optimal synthesis will result in a high proportion of short cDNAs, thereby reducing the amount of target DNA produced from the 5' end. Consequently the amplification of this region will be very difficult.

Gene specific primers were also tested for use in cDNA synthesis in an attempt to counteract the problems encountered with the oligo dT and random hexamer primers. The cDNA was synthesised in three overlapping pieces using the reverse primers for fragments II, IV and VI (see Figure 3.1). Amplification was carried out on the corresponding fragments and the fragment immediately upstream. With this method of synthesis it was not possible to check the quality of the cDNA by amplification of a control PCR fragment. Therefore the optimisation of cDNA synthesis and PCR had to be combined, which will be discussed further in section 3.4.

There are two advantages to the use of gene specific primers for cDNA synthesis. Firstly, cDNA is generated only from the target gene eliminating a dilution effect from non-specific sequences, greatly reducing the chances of mispriming in PCR and inhibition of amplification from too much DNA in the reaction. Secondly, several primers along the gene can be used in synthesis, reducing the extent of elongation required in each reaction. In this case, for *CSB*, each primer was only expected to generate a maximum of 3kb of cDNA.

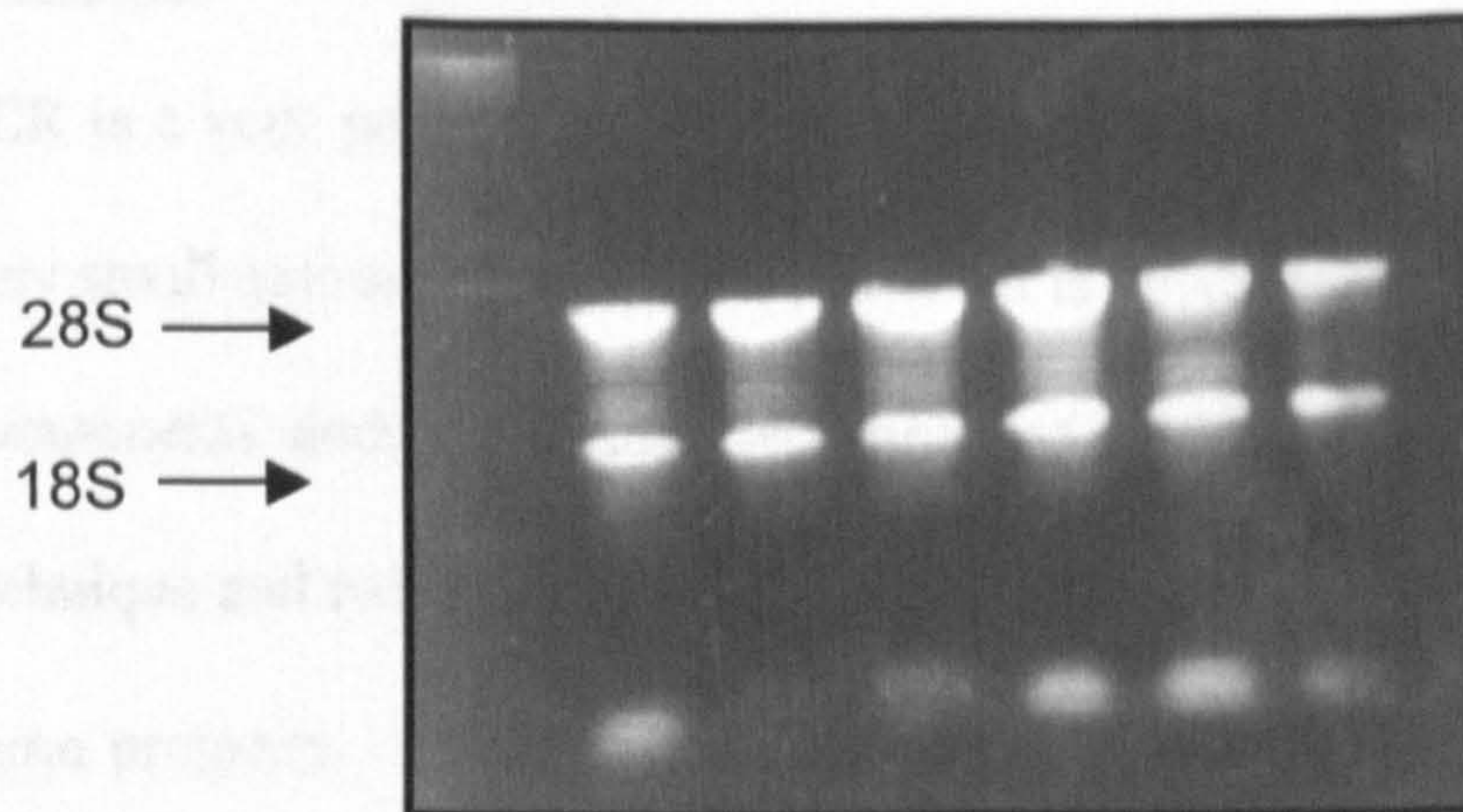


Figure 3.2 Formamide agarose gel of a standard total RNA preparation, demonstrating the relative positions of the 28S and 18S ribosomal RNA bands

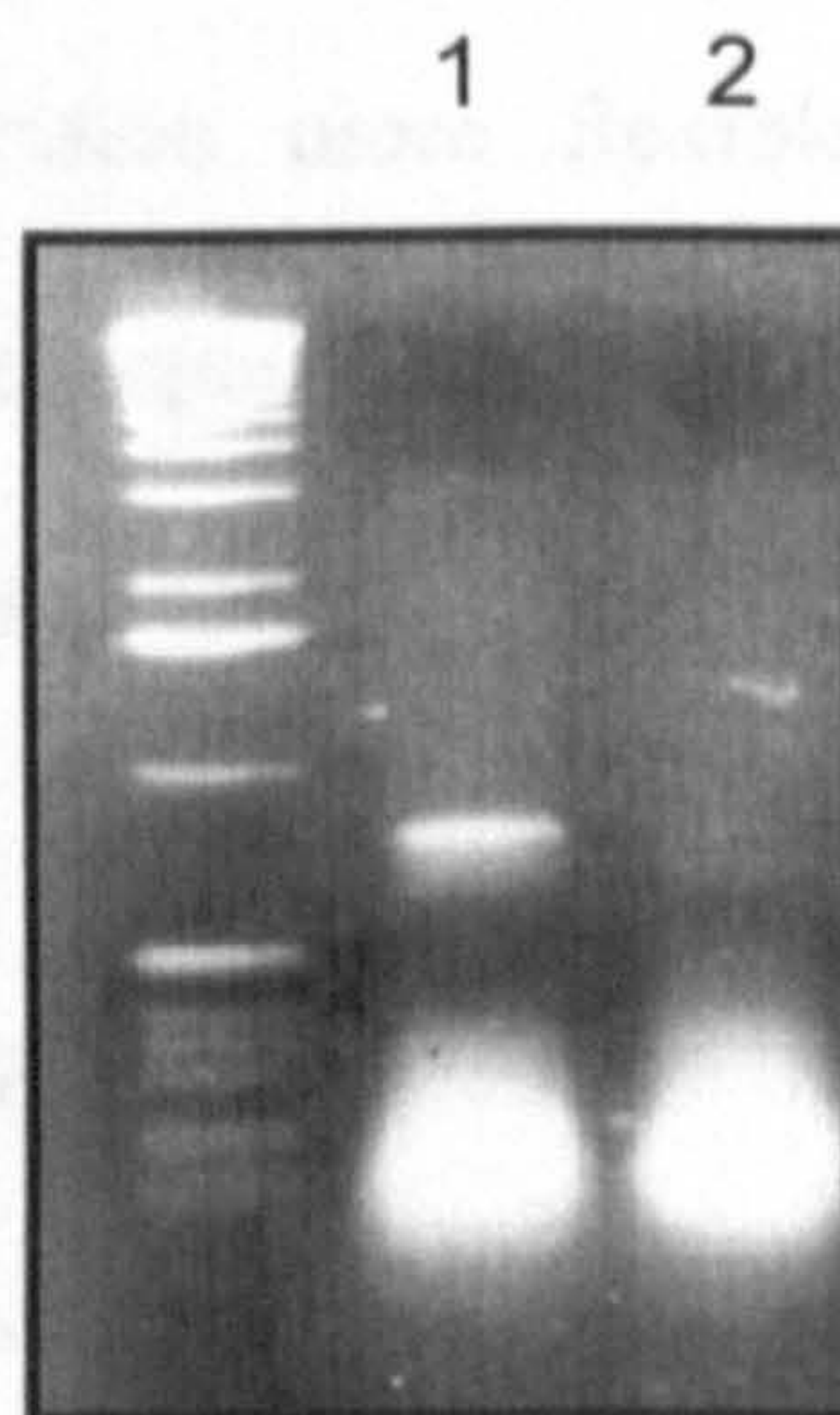


Figure 3.3 Agarose gel showing the control *hpert* PCR product amplified from cDNAs primed with oligo dT (lane 1) and random hexamers (lane 2).

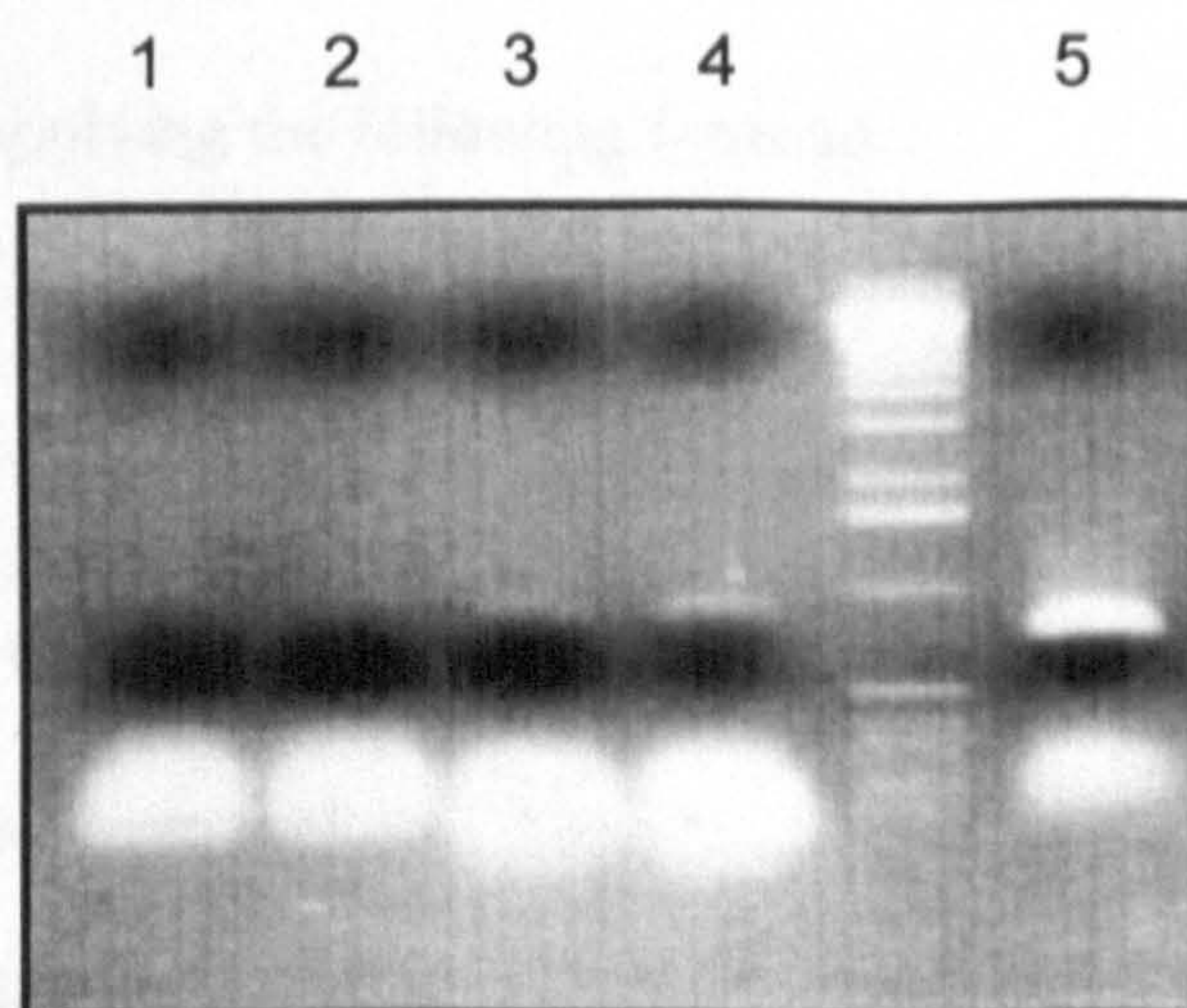


Figure 3.4 Agarose gel showing amplification of the 5' end of *CSB*, demonstrating the weak signal obtained for fragment II (lanes 3 and 4), and the barely visible product for fragment I (lanes 1 and 2). Lane 5 shows amplification of the control *hpert* fragment.

3.4 PCR

PCR is a very powerful technique allowing generation of large quantities of DNA from a very small amount of starting material. It is however, highly dependent on the nature of its components and the conditions employed during the reaction. PCR is also a variable technique and two thermal cyclers will rarely give identical conditions even when set to the same program. The starting material can also have a great influence on the conditions required for efficient amplification. For example, if the target DNA is plasmid there is a much greater amount of target material in the sample compared to that present when amplifying from an equivalent amount of cDNA. A greater amount of starting material will make the amplification more flexible, allowing amplification at non-optimal conditions. It is therefore important when trying to amplify from a low abundance cDNA that conditions are optimised carefully.

3.4.1 Optimisation of conditions

Conditions were optimised using cDNAs synthesised by each of the above methods for the six fragments used in the amplification of *CSB*. Initially, primers obtained from Rotterdam were tested at the conditions suggested, then checked with a series of different annealing temperatures. Annealing temperatures of oligonucleotides can also be predicted from their composition by applying the following formula:

$$4 (G + C) + 2 (A + T) - 5.$$

Each primer pair was checked at three annealing temperatures between 60°C and 70°C. As described below three primer sets did not give reliable products and therefore new primers were designed for these regions. The new primers were also checked to find the optimal annealing conditions.

3.4.2 Results and discussion

Most of the primer pairs gave products at one of the annealing temperatures tested. Fragment I however, gave unreliable PCR products unrelated to the quality of the cDNA, with the same cDNA unable to give reproducible amplification. One explanation for this erratic amplification could be a large amount of secondary structure, common in the 5' untranslated region of genes, preventing the primers from annealing to the target sequence. This problem was overcome by introducing a hot start step into the PCR. This involves heating the reaction to 94°C for 5 min before beginning the cycling reaction to eliminate secondary structure from the first round of PCR. Once this problem had been overcome, the amplification of this fragment could also be optimised.

The six fragments could be amplified with a wide range of annealing temperature, using a standard amplification protocol as outlined in Table 3.1 below.

Fragment	Hot Start	Denaturation	Annealing	Elongation	No. of cycles
I	94°C 5 min	94°C 1 min	65°C 1 min	72°C 2 min	35
II	94°C 5 min	94°C 1 min	63°C 1 min	72°C 2 min	35
III	94°C 5 min	94°C 1 min	67°C 1 min	72°C 2 min	35
IV	94°C 5 min	94°C 1 min	67°C 1 min	72°C 2 min	35
V	94°C 5 min	94°C 1 min	67°C 1 min	72°C 2 min	35
VI	94°C 5 min	94°C 1 min	60°C 1 min	72°C 2 min	35

Table 3.1 The PCR cycling conditions employed for the optimal amplification of the six fragments encompassing the whole ORF of the *CSB* cDNA.

3.5 Sequencing

Two approaches can be taken when sequencing PCR products, cloning and direct sequencing. The cloning procedure takes advantage of a property of the Taq Polymerase

which adds a single A nucleotide onto the 5' end of each strand of amplified DNA. The PCR products can therefore be cloned easily into a T-vector which has single T overhangs. The clones allow generation of ssDNA and production of large quantities of DNA. However, several clones from each PCR product must be sequenced to ensure that any changes identified are mutations rather than PCR errors. This method is particularly useful when studying patients that have compound heterozygous mutations or splice abnormalities. In such cases it is necessary to identify differences between alleles.

Alternatively, the PCR product can be sequenced directly, preventing PCR errors being visible on the sequencing gel and being confused with mutations. Direct sequencing also allows simple distinction between heterozygous and homozygous mutation, where in the case of homozygous changes only the mutant band is observed. However, when looking at sequence with a heterozygous change, both the mutant and parental bands will be present at the same position.

For the purposes of this study, cloning was used as a back-up procedure when direct sequencing failed to give sufficient information as to the nature of the mutation. Direct sequencing is a far more rapid and labour efficient way of sequencing PCR products. Therefore, it was decided to test two available techniques for direct sequencing for use in this project.

3.5.1 Cycle sequencing

Cycle sequencing is based on PCR and uses a modified Taq polymerase. The method works by incorporating radioactively labelled nucleotides into the DNA via a primer extension reaction then mixing with dideoxy terminators.

3.5.2 Strand separation

The strand separation technique works by tagging one strand of the DNA by re-amplifying the first round PCR product using a biotin tagged nested primer. The DNA is bound onto streptavidin coated magnetic beads via the biotin and the two strands are separated by alkaline denaturation with the tagged strand held with a magnet. The resulting single-stranded DNA can be sequenced as usual using the Amersham Sequenase 2.0 kit as described in Chapter 2.

3.5.3 Results and discussion

The two sequencing methods were initially tested for sequencing of *CSB* fragment VI which results in a pure, high yield PCR product. Both methods produced good results with clear sequencing ladders and a minimal amount of cross-banding. From these results it was difficult to determine the method most suitable for sequencing *CSB*, with the strength of signal the only variation, being slightly stronger with the strand separation method. Therefore, the other *CSB* PCR fragments were tested with both methods. When sequencing fragment V, the strand separation method again gave good clear results. However, for cycle sequencing clear sequence was obtained for very few of the reactions, most resulting in considerable cross-banding or smearing.

3.6 Conclusions

Many different techniques were tested with the *CSB* gene to determine the best methods and conditions for the amplification and sequencing of the gene in CS patients.

For RNA extraction both methods have been employed during the project, with the cell lysis procedure being preferred when genomic DNA is also required for analysis. For cDNA synthesis, both priming by oligo dT and specific primers gave amplification.

Because priming with the specific primers gave the more reliable results this method was used, despite it being necessary to make three cDNAs for each patient.

Strand separation was the method of choice for direct sequencing, giving much clearer and therefore more reliable data. More recently a new sequencing method has been used in the laboratory (Amersham). It is able to eliminate the cross-banding seen particularly in cycle sequencing, due to the use of ^{33}P -labelled ddNTP terminators. It is also more rapid and reliable than either of the sequencing methods described here, leading to the method being used routinely.

The techniques discussed here have been used for the analysis of the *CSB* gene in Cockayne Syndrome patients. The results of these analyses are shown in chapter 4.

CHAPTER 4 - MUTATIONS IN *CSB*

4.1 Introduction

Cockayne Syndrome (CS) cells are deficient in transcription coupled repair and fail to recover RNA synthesis after UV-C irradiation. Patients suffering from Cockayne Syndrome can therefore be diagnosed by testing for the ability of fibroblasts grown from a skin biopsy to recover RNA synthesis after irradiation. This is illustrated in Fig 4.1 which shows the recovery of RNA synthesis in normal and mutant cell lines at a range of UV doses.

There are two complementation groups of CS, A and B with the majority of patients belonging to group B (Stefanini et al., 1996). The genes defective in each of the complementation groups have been cloned, the *CSA* gene in 1995 by Henning et al., 1995 and *CSB* in 1990 by Troelstra et al., 1990. In this study the *CSB* gene was sequenced in patients from complementation group B, in an attempt to identify the causative mutations. It was hoped that such an analysis would identify regions of the gene essential for function and determine a link between the location of the mutation, the clinical severity and the nature of the CS symptoms.

As described in Chapter 1, *CSB* is a member of the SWI/SNF superfamily of putative helicases and ATPases. All members of this family contain the consensus sequences of the seven helicase domains which are highly conserved between species. If such regions are important for function, it can be predicted that mutations falling within this area may inactivate the protein.

Structural homologues of *CSB* have been found and sequenced in mouse and the yeast *Saccharomyces cerevisiae*. Mutations found in the cell lines studied can be compared with the mouse and yeast sequences to discover whether they fall within the conserved regions of the gene or change important residues. The nature of the amino acid change can also

give important clues as to whether the mutation is a polymorphism or an inactivating change. A large change in size, charge or hydrophilic/hydrophobic nature between the original residue and the substituted amino acid is likely to prove the most disruptive to the protein.

4.2 Materials and Methods

The techniques outlined in Chapter 3 were used for the sequencing of the *CSB* gene for patients with Cockayne Syndrome from complementation group B. In brief, the methodology consisted of total RNA extraction from primary fibroblast cell lines and cDNA was synthesised by reverse transcription primed in three fragments with the gene specific primers R2, R4 and R6 (see table 2.1). The cDNA was then subjected to PCR amplification under the conditions outlined in Table 3.1 and sequenced directly. The sense strand is biotin tagged and bound to streptavidin coated magnetic beads. The tagged strand can be separated from the complementary strand and sequenced using the Sequenase Version 2.0 kit (for more detail, see section 3.3). Mutations not resulting in deletions, frameshifts or terminations were created in a mammalian expression vector by site-directed mutagenesis and cloning. The subsequent mutant genes were transfected into the mutant hamster cell line UV61, which is in the CS-B complementation group, to check for the ability to correct the UV-C sensitive phenotype. Stably transfected clones were analysed for UV survival.

In the case of CS1MA, a region of the genomic DNA was sequenced to check for mutations at the intron/exon boundaries. Genomic DNA was purified from the supernatant of the alkaline lysis RNA preparation (see chapter 2). The supernatant was treated with RNase (1µg/ml final concentration) at 37°C for 2 hr, then proteinase K (1µg/ml final concentration) at 65°C for 1 hr. The solution was extracted once each with equal volumes of phenol, phenol:chloroform/iso-amyl alcohol (50:50) and chloroform/iso-amyl alcohol

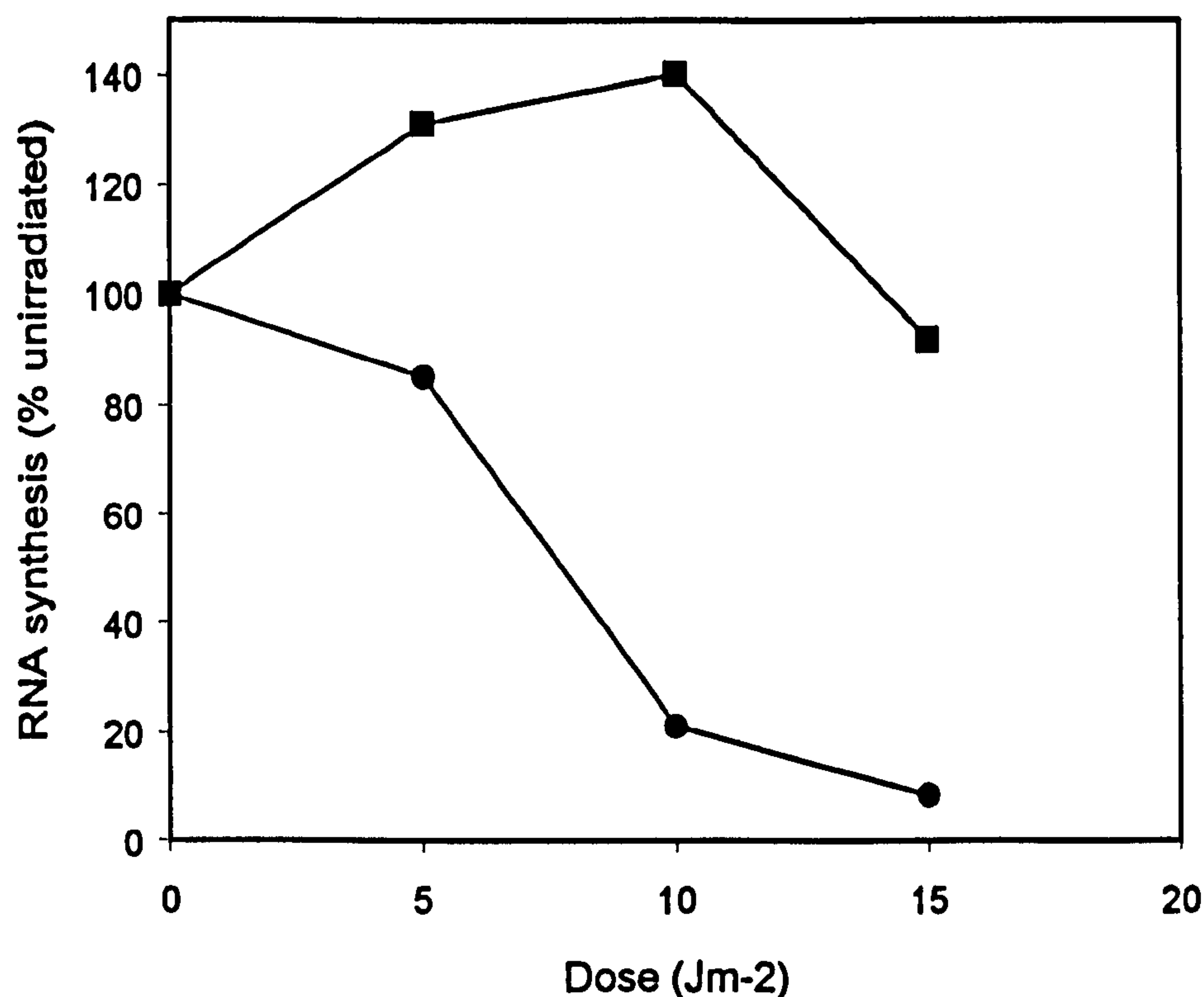


Figure 4.1 The incorporation of ^3H -uridine into CS-B (●) and wild-type (■) cells demonstrating the levels of RNA synthesis 24 hr after UV irradiation (data courtesy of H. Fawcett).

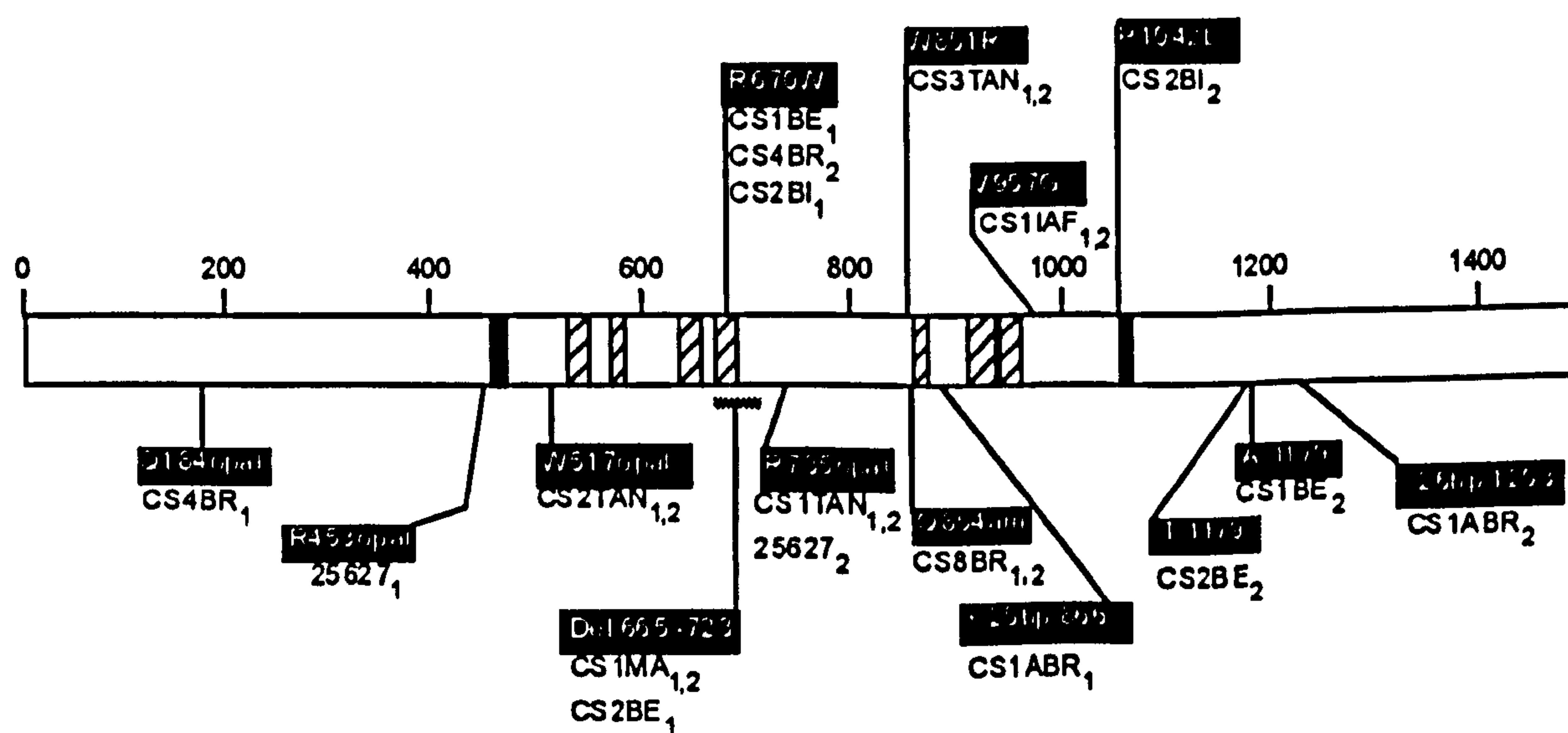


Figure 4.2 Diagrammatic representation of CSB illustrating the position and nature of mutations found in CS-B patients. The CSB protein is shown with the seven domains conserved in DNA helicases/ATPases(hatched boxes) and the putative nuclear-localisation signals(black boxes). Amino acid changes resulting from mutations are shown boxed with the change indicated by white lettering on black background, and the cell line designations are shown below. The subscripts 1 and 2 denote the different alleles. Mutations shown below the depicted protein result in protein truncations, whereas those above are single amino acid changes.

(24:1). Ethanol was added as 2½ volumes of the sample and the tube inverted twice to mix. The DNA was spooled from solution and resuspended in 500µl sterile double distilled water overnight with shaking at 37°C.

New primers were designed from intron sequence (C. Troelstra, Rotterdam) to allow amplification of the boundaries of introns 9 and 10 with exon 10 (see Appendix I for primer sequences).

4.3 Results

Thirteen patients were analysed that had previously been diagnosed as CS through failure to recover RNA synthesis after UV irradiation, and subsequently assigned to the CS-B group by somatic cell fusion and complementation assays (Lehmann, 1982; Stefanini et al., 1996). The *CSB* cDNA was sequenced in each case and compared to the published sequence. Two of these patients were siblings presenting a very unusual pattern of mutations, which due to their complex and interesting nature will be discussed in a later chapter.

In total, four types of mutations were found in the patients studied, as a result of either deletions or base substitutions. Two categories of deletion were observed in the *CSB* cDNA during this project. Two patients possess a deletion resulting from aberrant splicing, which has caused the exon to be removed, along with the flanking introns, possibly due to mutations at the splice junctions. Secondly, three patients carry single base deletions within a run of repeated nucleotides. This causes a frameshift downstream of the deletion, altering the amino acid sequence such that there is premature termination of the transcript and truncated protein is produced.

Base substitutions can give rise to two types of mutation, amino acid substitutions and termination codons. In the case of the latter, the effect of replacing an amino acid with a stop codon results clearly in truncated protein. Amino acid substitutions are far more

Cell line ^a	aa change ^b	Racial Group	Age ^c	Age of Onset	Pigment. Retinop. ^d	Gait ^e	Cataracts	Caries	Ref
CS2TAN (88) YK	Trp517stop	Turkish	9	1.5	y	y	n	n	88 in [Lehmann, 1993 #604]
CS1MA (34) LW	Del 665-723	Caucasian	4 dead	0	?	y	y	y	34 in [Lehmann, 1993 #604]
CS1ABR (107) RC	+ 25 866 + 26 1203	Caucasian	6	0.2			y		[Stefanini, 1996 #5959]
CS1TAN (77) YA	Arg735stop	Turkish	8	1.5	y	y	n	y	77 in [Lehmann, 1993 #604]
25627 JM	Arg735stop Arg453stop	Caucasian	5		y	y			
CS2BE	-1 1179 Del 665-723	Caucasian	21	1	y	y			[Brumback, 1978 #1094]
CS4BR (39)	Arg670Trp Gln184stop	Caucasian	24	3	y	y	y	y	39 in [Lehmann, 1993 #604]
CS2BI (9) MS	Arg670Trp Pro1042Leu	Caucasian	16	0.7	y	y			9 in [Lehmann, 1993 #604]
CS1BE	Arg670Trp -1 1179	Caucasian	11	1.5	y	y	n		[Brumback, 1978 #1094]
CS3TAN (98) MI	Trp851Arg	Turkish	1.7	0.3	n	y	n	n	98 in [Lehmann, 1993 #604]
CS1IAF (67) IN	Val957Gly	Israeli Jewish	7	2	y			y	67 in [Lehmann, 1993 #604]

^a**Bold:** Patient with early onset form of disorder. ^b**Bold:** Less severe mutation. For compound heterozygotes, this is assumed to be the more C-terminal of two truncating mutations, or a mis-sense mutation if one allele is a truncation and the other a mis-sense. ^cAge at biopsy. ^dPigmentary retinopathy. ^eGait defects. y, yes; n, no.

Table 4.1 The origin, clinical features and mutations of the patients studied in this chapter.

complex, with the effect of the change not always obvious. It is therefore particularly important in these cases to determine whether the change is phenotypically significant. Initially, when mutations were found that resulted in amino acid substitutions, the entire gene was sequenced to check for the absence of any other changes. Subsequently, the cDNA was sequenced within the central fragments and not continued outside fragments II to V if putative inactivating mutations were found.

All mutations found during this section of the project are shown in Figure 4.2. This figure is a diagrammatic representation of CSB showing the position of putative nuclear localisation signals and the seven helicase domains. The full sequence of the CSB cDNA is shown in Appendix III.

4.3.1 CS2BE and CS1MA

CS1MA is the offspring of unrelated parents of Western origin and was therefore not expected to carry a homozygous mutation. However, there was an unusually small PCR product on amplification of fragment 3 for CS1MA, which on direct sequencing showed that amino acids 665 to 723 were absent (Fig 4.3). The deletion spans exon 10 which comprises 174 nucleotides and is positioned within helicase/ATPase domain III. The deletion is likely to be the result of a splicing abnormality. In an attempt to find the origins of the probable aberrant splicing, the boundaries between exon 10 and introns 9 and 10 were analysed in CS1MA, through PCR and sequencing of genomic DNA from these regions. The sequencing failed to show any mutations which may be responsible for the removal of exon 10 during splicing. The cause of this deletion remains unsolved.

CS2BE originates from the USA and has the typical features of a CS patient as shown in Table 4.1. This patient is a compound heterozygote, carrying a different mutation in each allele of *CSB*. The first allele has a deletion of exon 10 as described above for CS1MA that was noticed by the presence of two distinct bands on amplification of PCR fragment

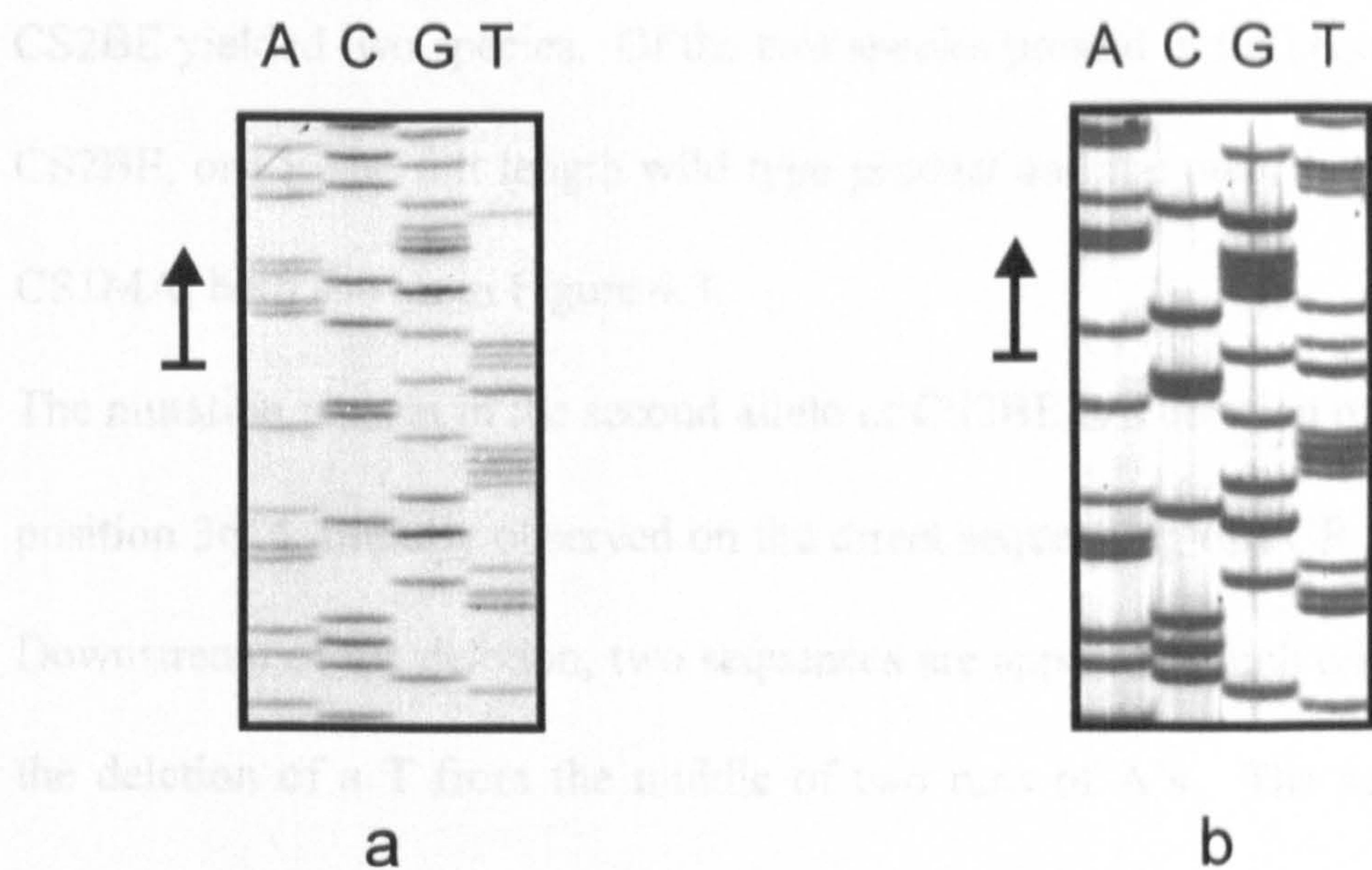


Figure 4.3 Sequence for patient CS1MA (panel a) and wild-type (panel b) around the boundary of exons 10 and 11, demonstrating the absence of exon 10 in CS1MA.

4.3.1 CS1A9H

CS1A9H is a Caucasian patient from Australia with the severe early onset form of Cockayne Syndrome (see Table 4.1) and is the offspring of a consanguineous marriage.

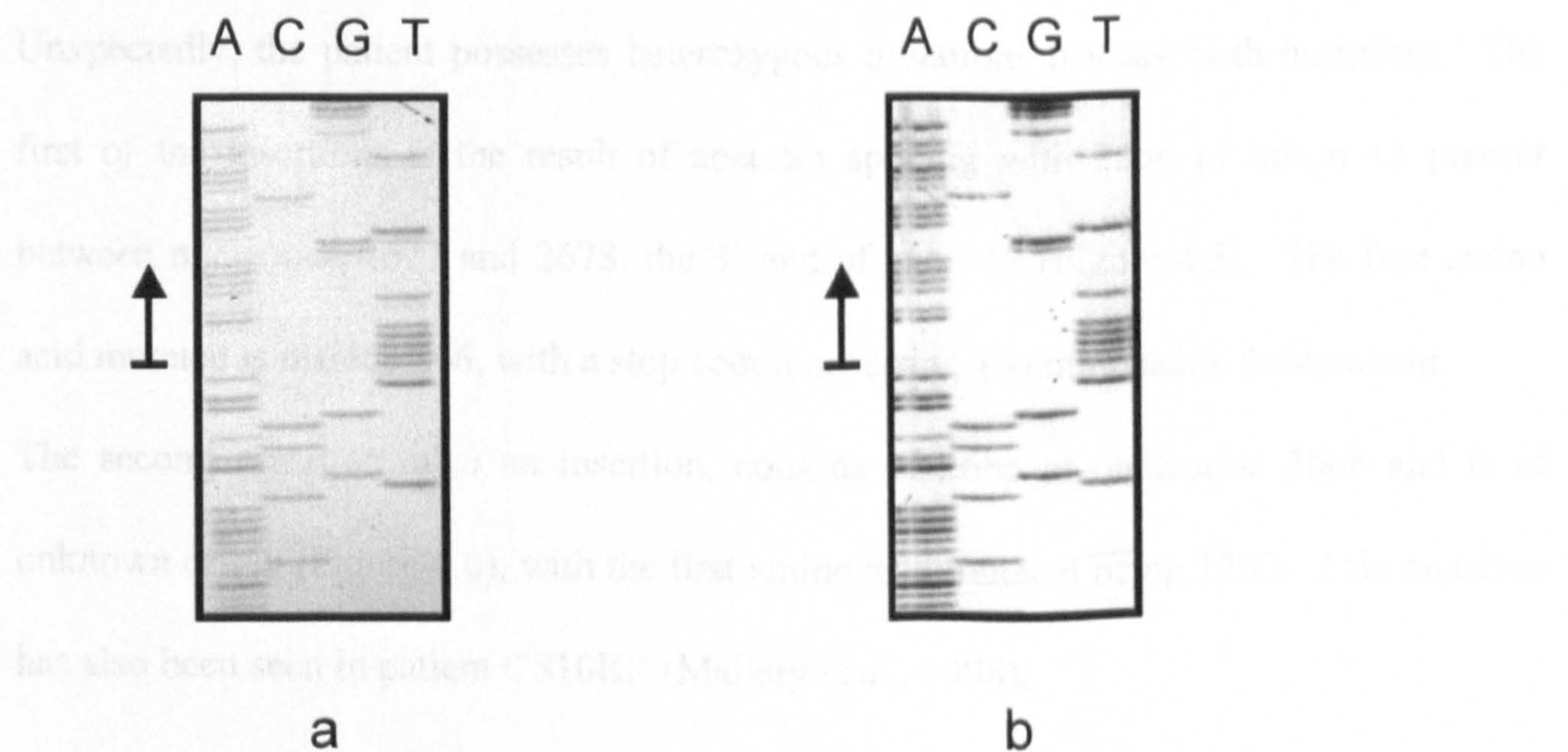


Figure 4.4 Sequence for patient CS2BE (panel a) and wild-type (panel b) around position 3614, showing the presence of both the wild-type and single T deleted alleles from the arrow upwards in a.

The two Turkish patients (CS1A7N and CS1A7O) are the offspring of consanguineous marriages and consequently both carry identical mutations in CS1A. The first patient, CS1A7N, possesses a C to T transition at position 2252 (see Figure 4.7).

III from CS2BE. These two products are also apparent on the sequencing gel where CS2BE yielded two species. Of the two species present in the sequencing of fragment 3 in CS2BE, one is the full length wild type product and the other is the deleted product from CS1MA, both shown in Figure 4.3.

The mutation present in the second allele of CS2BE is a deletion of a single T nucleotide at position 3614, initially observed on the direct sequencing of PCR fragment 5 (Figure 4.4). Downstream of the deletion, two sequences are apparent which on closer inspection reveal the deletion of a T from the middle of two runs of A's. The sequence of this allele is shifted by one base giving the shadowing effect seen in the panel a of Figure 4.4 when compared to the wild-type in the panel b.

4.3.2 CS1ABR

CS1ABR is a Caucasian patient from Australia with the severe early onset form of Cockayne Syndrome (see Table 4.1) and is the offspring of a consanguineous marriage. Unexpectedly, the patient possesses heterozygous mutations that are both insertions. The first of the insertions is the result of aberrant splicing with 25bp of intron 13 present between nucleotide 2677 and 2678, the 3' end of exon 13 (Figure 4.5). The first amino acid mutated is residue 866, with a stop codon occurring 10 amino acids downstream. The second mutation, also an insertion, consists of 26bp at nucleotide 3686 and is of unknown origin (Figure 4.6), with the first amino acid mutated being 1203. This mutation has also been seen in patient CS10BR (Mallery et al., 1998).

4.3.3 CS1TAN and CS2TAN

The two Turkish patients CS1TAN and CS2TAN are the offspring of consanguineous marriages and consequently both carry inactivating mutations in both alleles of the *CSB* gene. CS1TAN possesses a C to T transition at nucleotide 2282 illustrated in Figure 4.7b

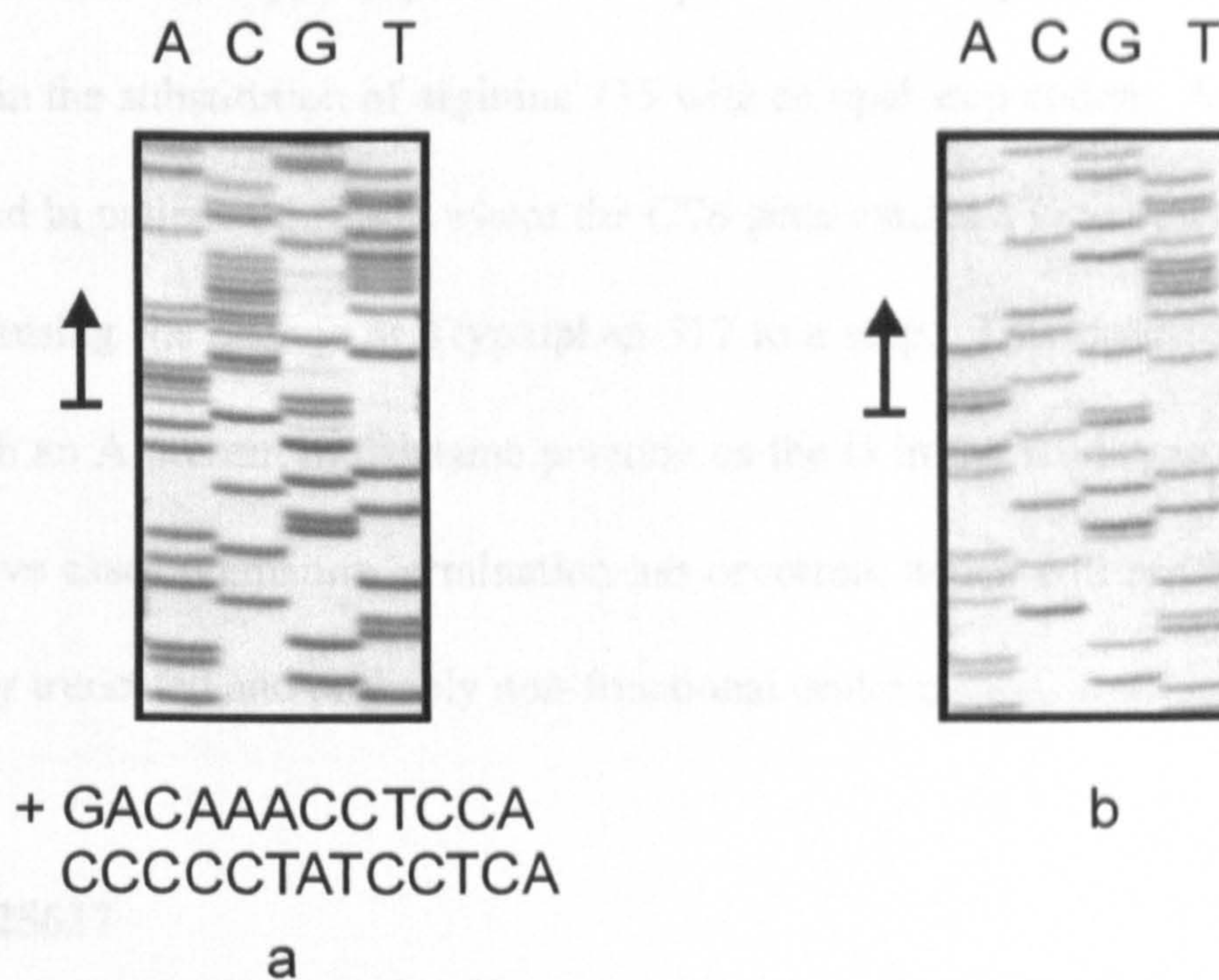


Figure 4.5 Sequence for CS1ABR (panel a) and wild-type (panel b) around the boundary of exons 13 and 14. Panel a shows the presence of both the wild-type sequence from panel b and the 25bp insertion from intron 13. The gels and sequence read in the 3' to 5' orientation.

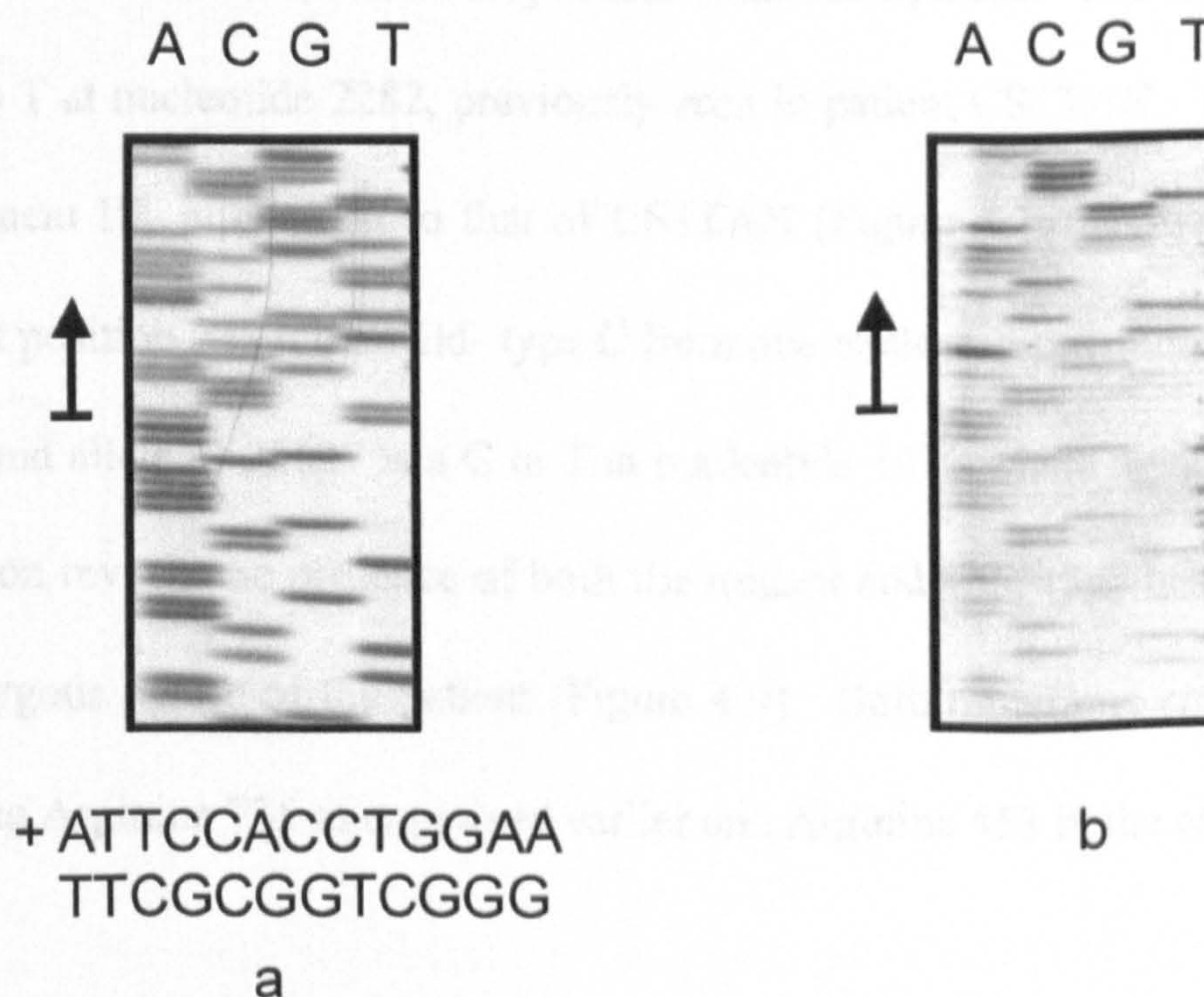


Figure 4.6 Sequence for CS1ABR (panel a) and wild-type (panel b) around position 3686. Panel a shows the presence of both the wild-type sequence shown in panel b and the 26bp insertion. The gels and sequence read in the 3' to 5' orientation.

along with the wild-type sequence at this position for comparison. This nucleotide change results in the substitution of arginine 735 with an opal stop codon. A similar occurrence is observed in patient CS2TAN where the *CSB* gene carries a G to A transversion at position 1630 causing the change of Tryptophan 517 to a stop. This mutation is shown in Figure 4.8 with an A present in the same position as the G in the wild-type sequence. In both of the above cases premature termination has occurred, which will result in the production of severely truncated and probably non-functional protein.

4.3.4 25627

The patient 25627 is the sibling of the sun-sensitive patient 11961 described by Arlett et. al., 1978 and assigned as CS group B by Lehmann, 1982. This patient has mutations in both alleles of *CSB* resulting in premature termination of the transcript, as discussed in the above section. However, in this case, the patient is a compound heterozygote with a different mutation in the allele originating from each parent. The first of the mutations is the C to T at nucleotide 2282, previously seen in patient CS1TAN. The direct sequencing of fragment III, in contrast to that of CS1TAN (Figure 4.7a) shows the presence of two bands at position 2282, the wild-type C from one allele and the mutant T. The mutation in the second allele of 25627 is a C to T at nucleotide 1436 where again direct sequencing of the region reveals the presence of both the mutant and wild-type bases, demonstrating the heterozygous nature of the patient (Figure 4.9). Both mutations create opal stop codons replacing Arginine 735 as explained earlier and Arginine 453 in the case of C1436T.

4.3.5 CS3TAN and CS1IAF

Patient CS3TAN is a Turkish patient and CS1IAF is of Israeli-Jewish origin. Both patients are the offspring of consanguineous marriages and consequently possess homozygous mutations. In contrast to the mutations discussed earlier, which all result in truncated

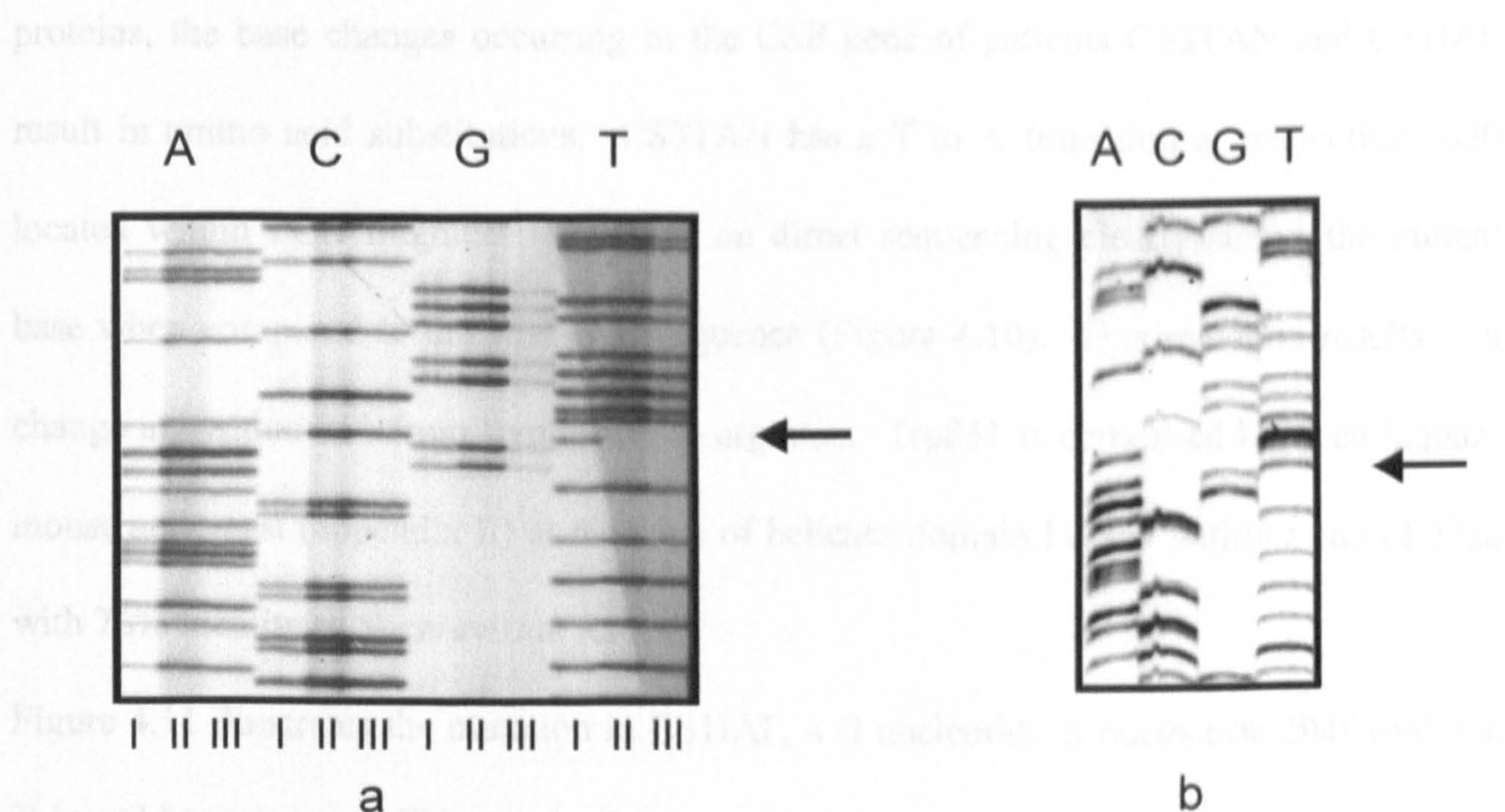


Figure 4.7 Sequence of patient 25627 and wild-type (panel a) showing the presence of bands in lanes I in both the T and C tracks around position 2282. Panel b shows the same C to T change this time homozygous, present in patient CS1TAN.

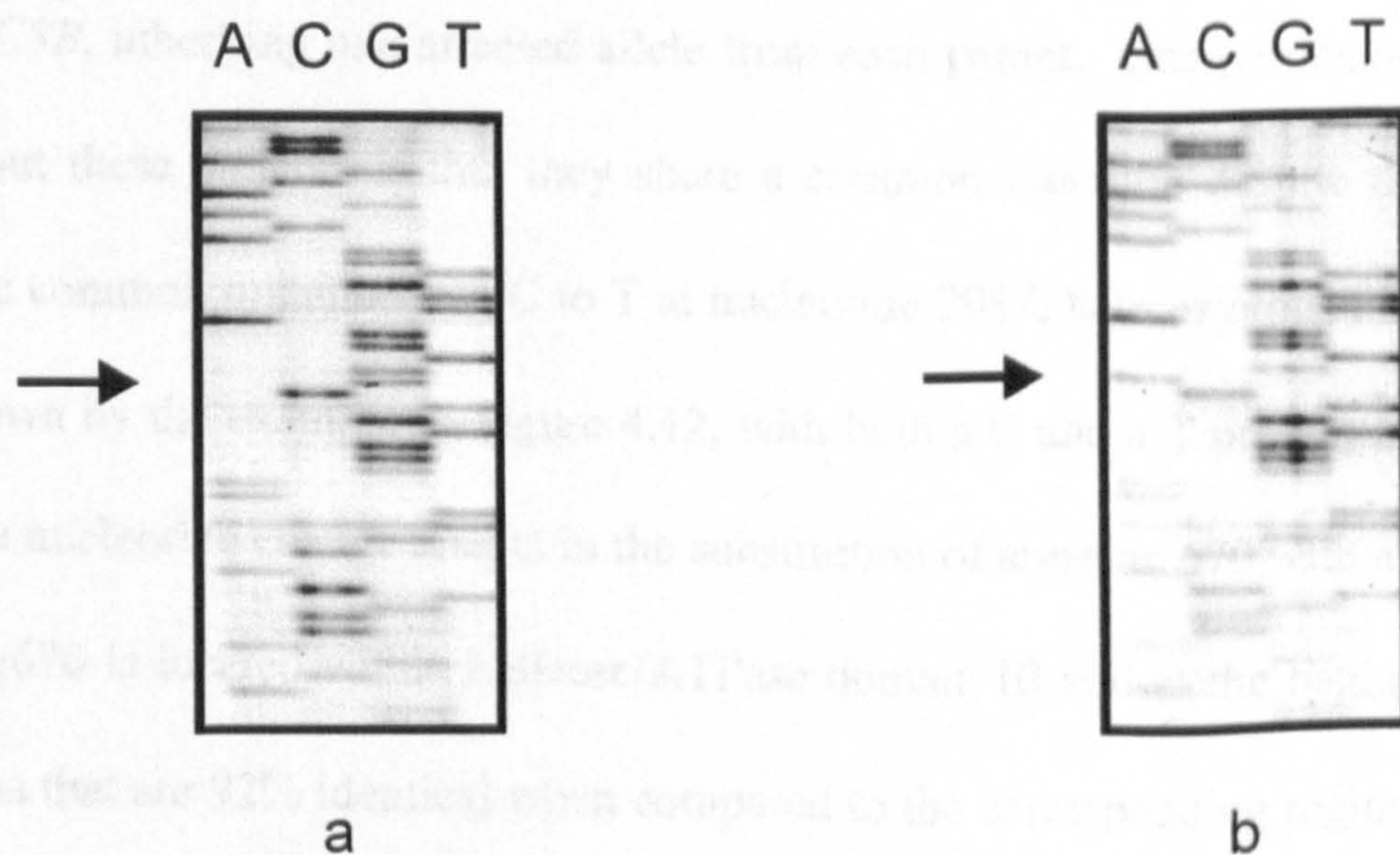


Figure 4.8 Sequence of wild-type (panel a) and patient CS2TAN (panel b) around position 1630 showing a homozygous G to A change in the patient.

proteins, the base changes occurring in the *CSB* gene of patients CS3TAN and CS1IAF result in amino acid substitutions. CS3TAN has a T to A transition at nucleotide 2630 located within PCR fragment IV, which on direct sequencing clearly shows the mutant base when compared to the wild-type sequence (Figure 4.10). This mutation results in a change at residue 851 from tryptophan to arginine. Trp851 is conserved between human, mouse and yeast (appendix II) at the edge of helicase domain IV and within a run of 37aa with 76% identity to *S. cerevisiae* Rad26.

Figure 4.11 illustrates the mutation in CS1IAF, a G nucleotide at nucleotide 2949 replaces T in wild-type cases. This results in a valine to glycine substitution at residue 957. As with the W851R mutation in CS3TAN residue 957 is highly conserved and is situated in a stretch of 79 aa with a yeast:human identity level of 83%.

4.3.6 CS1BE, CS4BR and CS2BI

The three patients CS1BE, CS4BR and CS2BI are compound heterozygous for mutations in *CSB*, inheriting one affected allele from each parent. One particularly interesting point about these patients is that they share a common mutation despite their diverse origins. The common mutation is a C to T at nucleotide 2087, heterozygous in all three patients as shown by the example in Figure 4.12, with both a C and a T present in the same position. The nucleotide change results in the substitution of arginine 670 with a tryptophan residue. Arg670 is located within helicase/ATPase domain III and at the beginning of a stretch of 34aa that are 92% identical when compared to the corresponding region in the *S.cerevisiae* homologue Rad26 (van Gool et al., 1994).

The second alleles in each case carry different mutations. CS1BE has a deletion of a single A at 3615 from within a run of four identical nucleotides and is illustrated in Figure 4.13. This is manifest by the sequence of the mutant allele in this position being offset by one base relative to the WT sequence. The deletion of A at 3615 causes a frameshift at amino

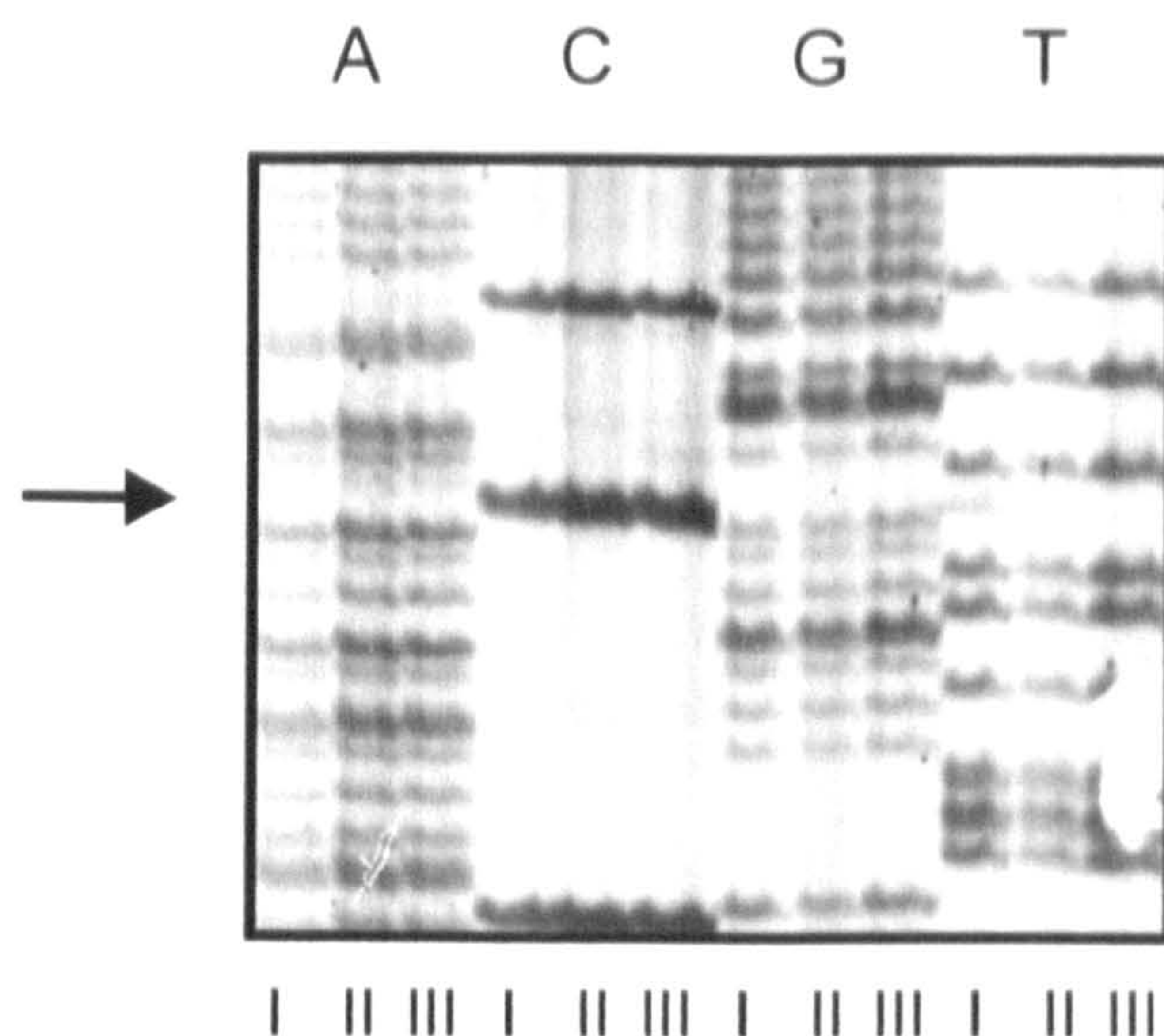


Figure 4.9 Sequence of patient 25627 (lanes I) and wild-type (lanes II and III) around position 1436, showing the presence of both nucleotides C and T in the patient lane.

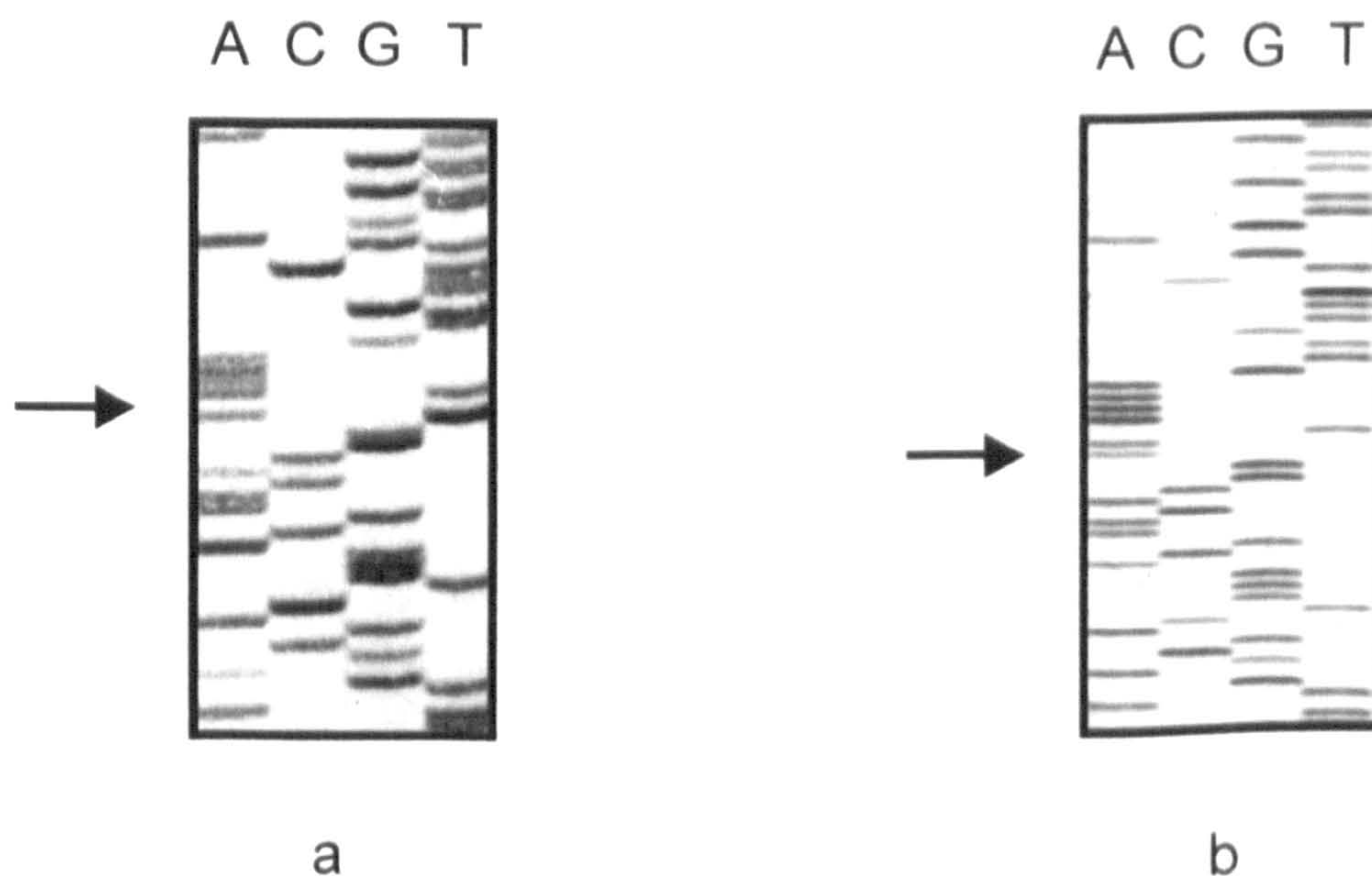


Figure 4.10 Sequence of wild-type (panel a) and patient CS3TAN (panel b) at position 2630 showing an A in place of T in panel b.

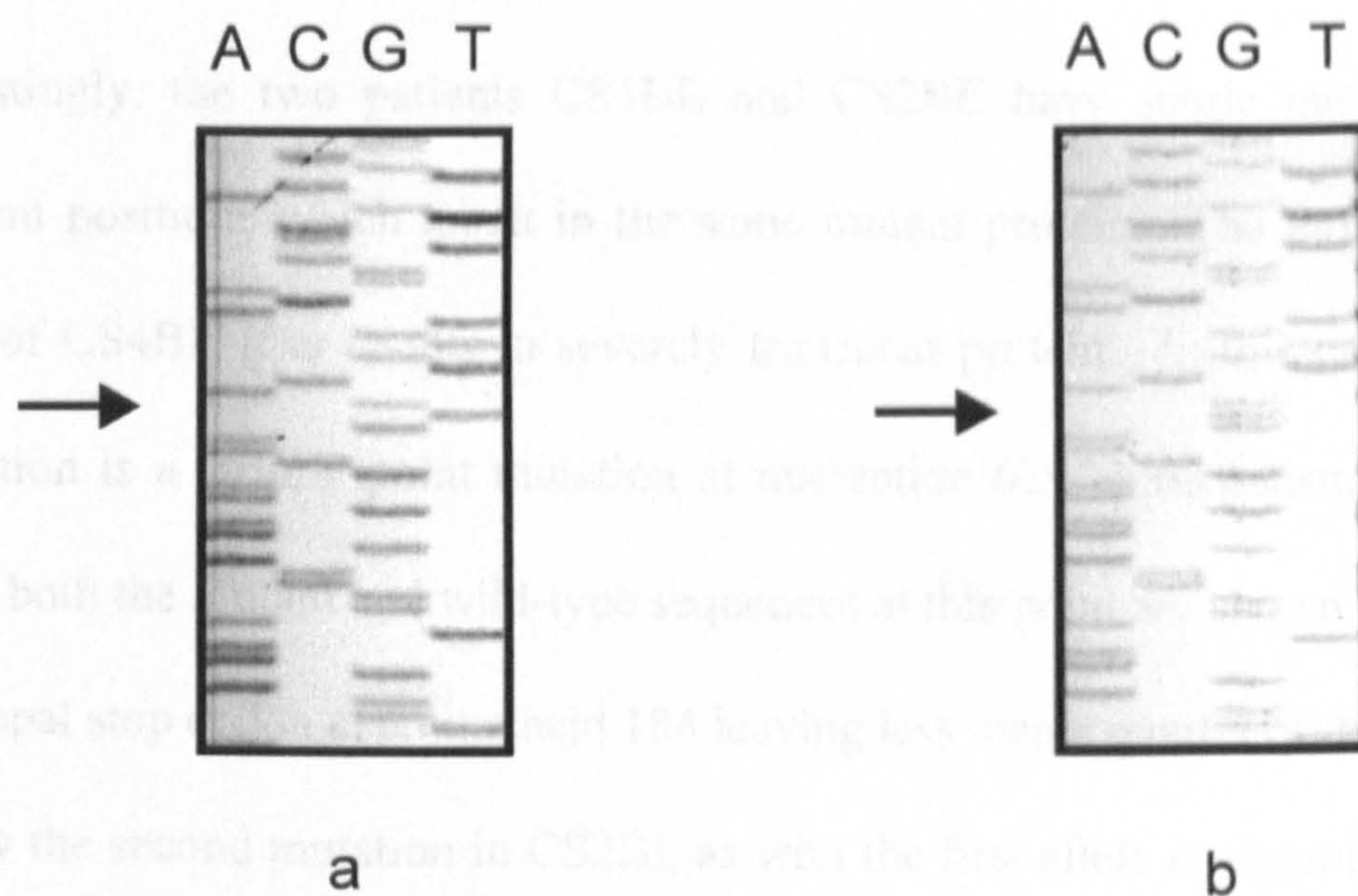


Figure 4.11 Sequence of wild-type (panel a) and CS1IAF (panel b) around position 2949 showing the nucleotide change from T to G in CS1IAF.

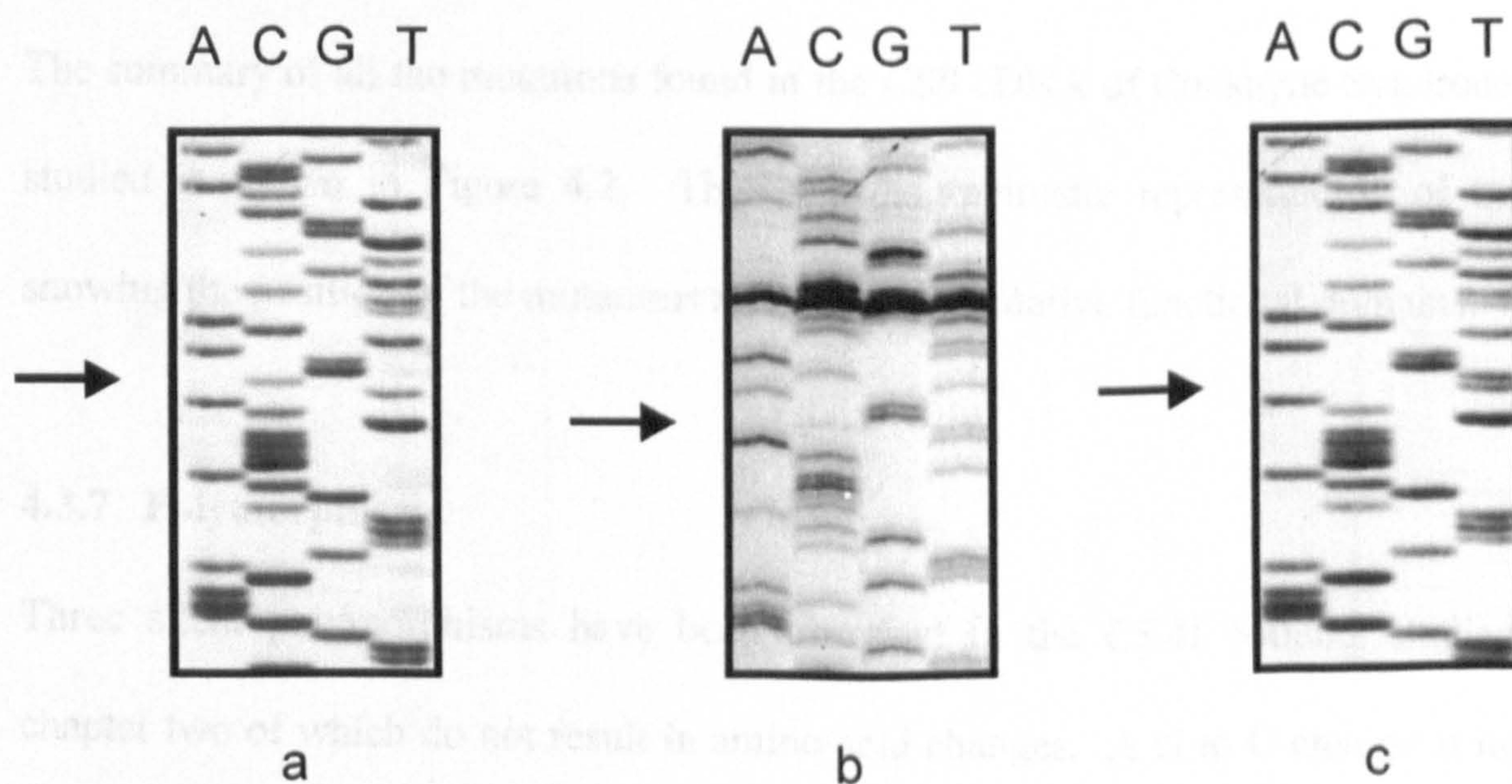


Figure 4.12 Sequence of wild-type (panel a), patient CS1BE, (panel b) and a homozygous mutant control (panel c) around position 2087. Panel b shows the presence of both the wild-type C and mutant T nucleotides at the same position.

acid 1179 and ultimately premature termination at residue 1200 as seen previously in patient CS2BE.

Interestingly, the two patients CS1BE and CS2BE have single nucleotide deletions at different positions which result in the same mutant protein. The mutation in the second allele of CS4BR also results in severely truncated protein. In this case, the cause of the truncation is a C to T point mutation at nucleotide 629 as demonstrated in Figure 4.14, where both the mutant and wild-type sequences at this point are shown. The change results in an opal stop codon at amino acid 184 leaving less than a quarter of the gene translated. Finally the second mutation in CS2BI, as with the first allele is an amino acid substitution mutation. This, as shown in Figure 4.15 is a T nucleotide at position 3204 instead of the wild-type C, with both bands being present on the sequencing gel in CS2BI. The mutation results in a proline to leucine change at aa1042. This change is not situated at a particularly highly conserved site, but is within the putative nuclear localisation signal, which is likely to be functionally significant.

The summary of all the mutations found in the *CSB* cDNA of Cockayne Syndrome patients studied is shown in Figure 4.2. This is a diagrammatic representation of the cDNA showing the position of the mutations relative to the putative functional domains.

4.3.7 Polymorphisms

Three silent polymorphisms have been observed in the CS-B patients studied in this chapter two of which do not result in amino acid changes. A G to C change at nucleotide 214 has been seen in the two patients CS2TAN and CS3TAN, and is located at leucine 45. It is possible that other patients within this study possess this polymorphism as this region was not sequenced in all the patients. Interestingly, the polymorphism at leucine 45 is the same amino acid as that mutated to proline in patient CS2GO (see chapter 5).

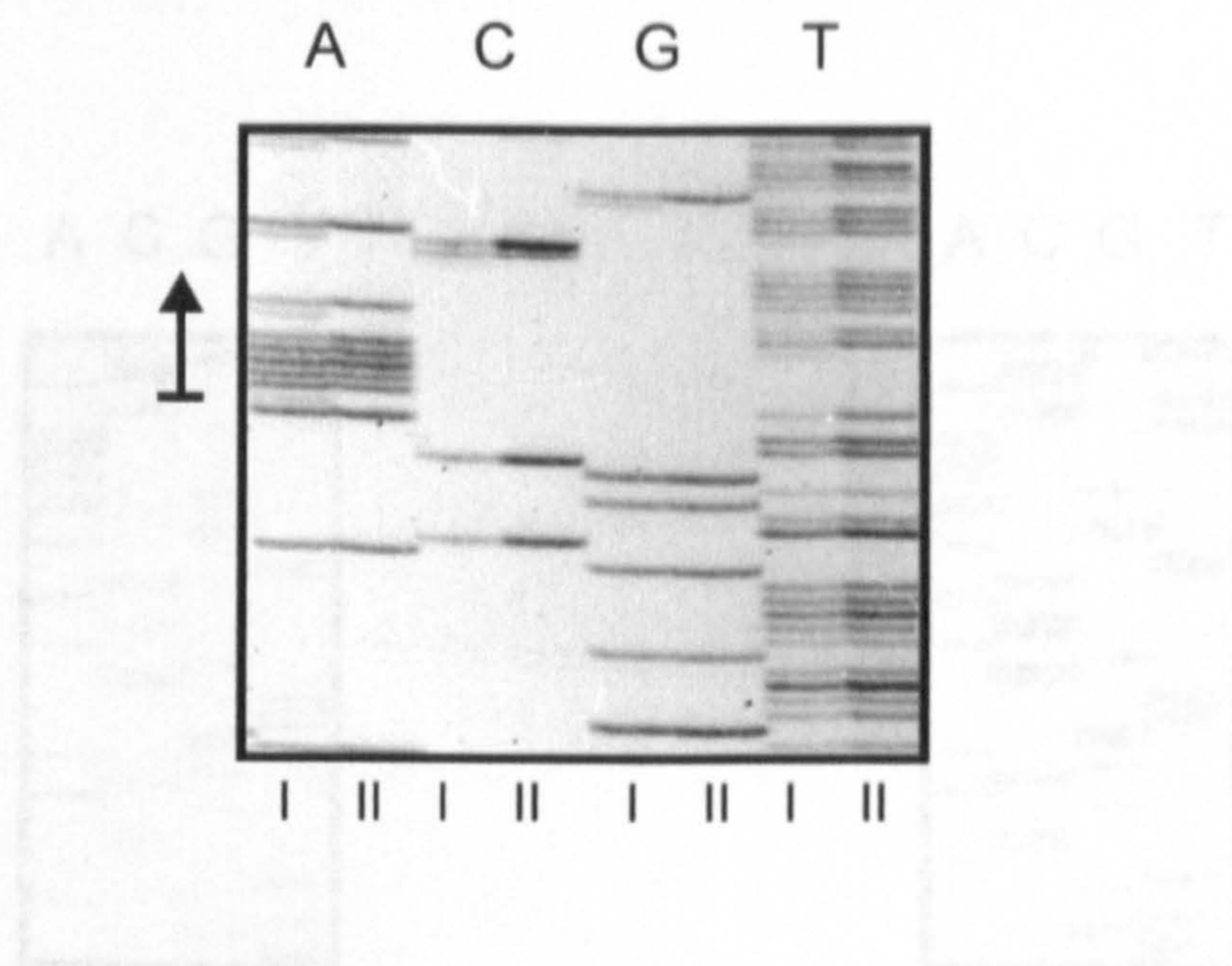


Figure 4.13 Sequence of wild-type (lanes II) and CS2BI (lanes I) around position 3615 showing that lanes I possess both the wild-type and single A deleted alleles.

Figure 4.13 Sequence of wild-type (panel a) and CS2BI (panel b) around position 3615 showing both the wild-type C and the mutant T present at the same position in panel b.

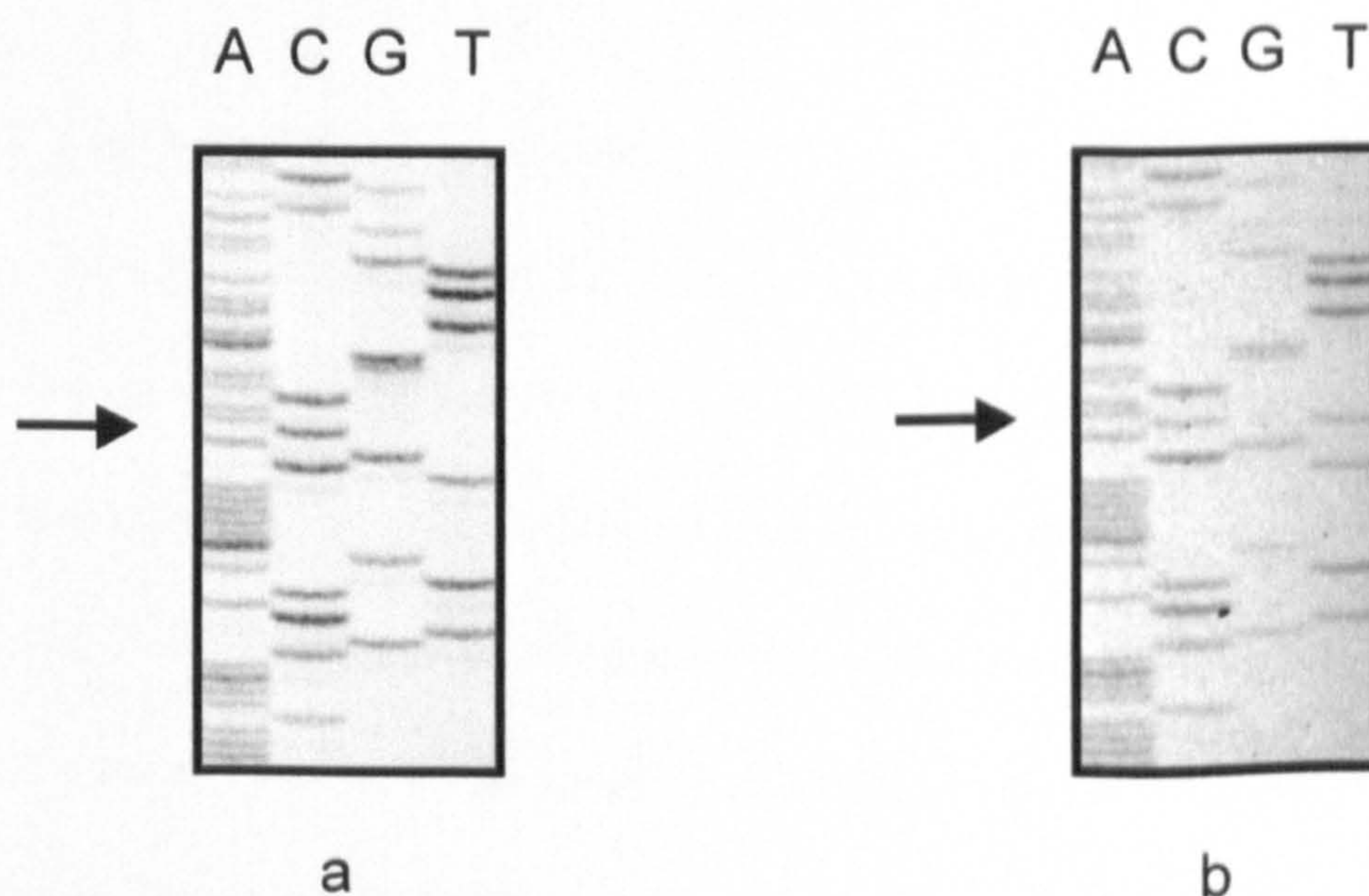


Figure 4.14 Sequence of wild-type (panel a) and CS4BR (panel b) around position 629 showing both the wild-type C and the mutant T are present in the same position in panel b.

The second polymorphism detected was a C to T change, which is associated with the CS2BI allele. This change has been checked in patients with CS2BI and in controls. The change has been checked in patients with CS2BI and in controls. The change has been checked in patients with CS2BI and in controls.

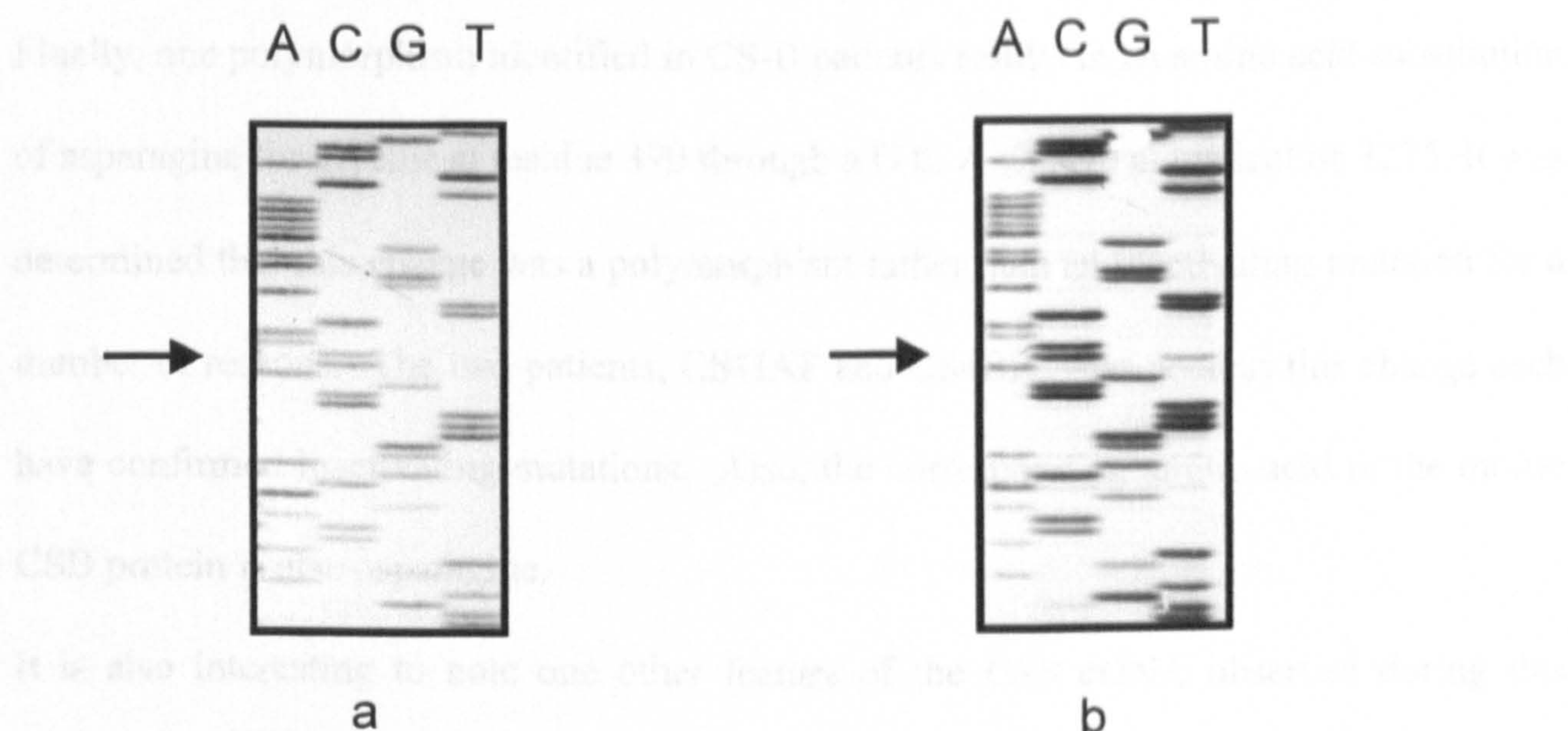


Figure 4.15 Sequence of wild-type (panel a) and CS2BI (panel b) around position 3204 showing both the wild-type C and the mutant T present at the same position in panel b.

It is likely that the consistent detection of a C to T change is due to a change in the sequence of the DNA.

4.4 Functional Complementations

As discussed above, many of the mutants found in CS2BI patients are predicted to be non-functional. To determine if these mutants are indeed non-functional, we performed functional complementation assays. In these assays, the mutant cDNA is co-transfected with a wild-type cDNA into cells. If the mutant cDNA is functional, it will complement the wild-type cDNA, and the cells will produce a functional protein. In this case, the mutant cDNA is not functional, and the cells will not produce a functional protein. This assay was performed for several mutants, and the results are shown in Table 1.

The results of the functional complementation assays show that the mutants found in CS2BI patients are indeed non-functional. This is consistent with the prediction that these mutants are non-functional.

The second polymorphism detected was a C to T change located at nucleotide 2830 and the amino acid 917. This change has been checked in all patients here and has been seen in three patients, CS1TAN, CS1BE and CS2BI.

Finally, one polymorphism identified in CS-B patients results in an amino acid substitution of asparagine for glycine at residue 399 through a G to A change at nucleotide 1275. It was determined that this change was a polymorphism rather than an inactivating mutation for a number of reasons. The two patients, CS1IAF and CS4BR, who possess this change each have confirmed inactivating mutations. Also, the corresponding amino acid in the mouse CSB protein is also asparagine.

It is also interesting to note one other feature of the *CSB* cDNA observed during this sequencing project. In all cases it was seen that a small proportion of the PCR fragment III products were missing exon 8 as determined by sequencing analysis. This is seen in the sequencing of all CS-B patients in this study, and in a cDNA sample wild-type for *CSB*. It is likely that the consistent deletion of exon 8 is the result of alternative splicing events.

4.4 Functional Complementation

As discussed above, many of the mutations found in CS-B patients result in substantial truncations, which almost certainly cause a total loss of function of CSB protein and are therefore null alleles. The mutations that fall into this category are the exon deletions, single base deletions and the resulting frameshifts and point mutations giving rise to stop codons. Point mutations that result in amino acid substitutions cannot with any certainty be said to result in non-functional protein. To be certain that amino acid changes found were inactivating mutations and not polymorphisms, they were tested for functional complementation in the hamster mutant cell line UV61.

The mis-sense mutations found in Cockayne syndrome patients were introduced into the *CSB* cDNA in pBluescript (generously provided by A. van Gool, Rotterdam) either by

subcloning of fragments used in sequence analysis or by site-directed mutagenesis. All PCR and mutagenised products were sequenced in entirety to ensure there were no errors. The resulting mutant genes were transferred from pBluescript into either of the mammalian expression vectors, pCDNA3 as a BamHI fragment in the earlier experiments, or pCI-neo as a XbaI fragment in later experiments.

The *CSB* cDNA was then transfected into the UV sensitive hamster cell line UV61 from the same complementation group as CS-B. Transfectants were selected by their ability to grow in medium containing the antibiotic G418. Resistance to G418 is conferred by the neo marker carried on the expression vector. Transfected cells were subsequently analysed both as a population and as single clones. The population was plated onto three 10cm dishes, two were irradiated with three doses of 4 Jm⁻² with a one day recovery period between each dose and the third was mock irradiated. Several days after the final dose the un-irradiated cells formed a confluent monolayer in all experiments. Similarly, irradiated UV61 cells transfected with wild type *CSB* cDNA grew to confluence within a few days. In contrast, irradiated cultures from cells containing the mutant cDNAs R670W, W851R, V957G and P1042L failed to show any significant cell growth. This is a strong indication that the mutations present result in a protein that is unable to correct the UV sensitive phenotype of UV61. Clones showing G418 resistance were picked and expanded for further analysis. These clones were checked for the presence of the *CSB* cDNA by PCR amplification from total genomic DNA using all six sets of primers. Two clones for each transfected cDNA that were shown to contain the full length cDNA were tested for UV survival. Survivals were carried out as a colony forming assay, using four different doses of 0, 3, 6 and 9 Jm⁻². The level of survival is shown as the number of colonies on irradiated plates expressed as a percentage of those on un-irradiated dishes.

Figure 4.16 shows the survival curves for the transfection of cDNAs expressed in pCDNA3. The wild-type gene when carried in this vector does not restore the survival

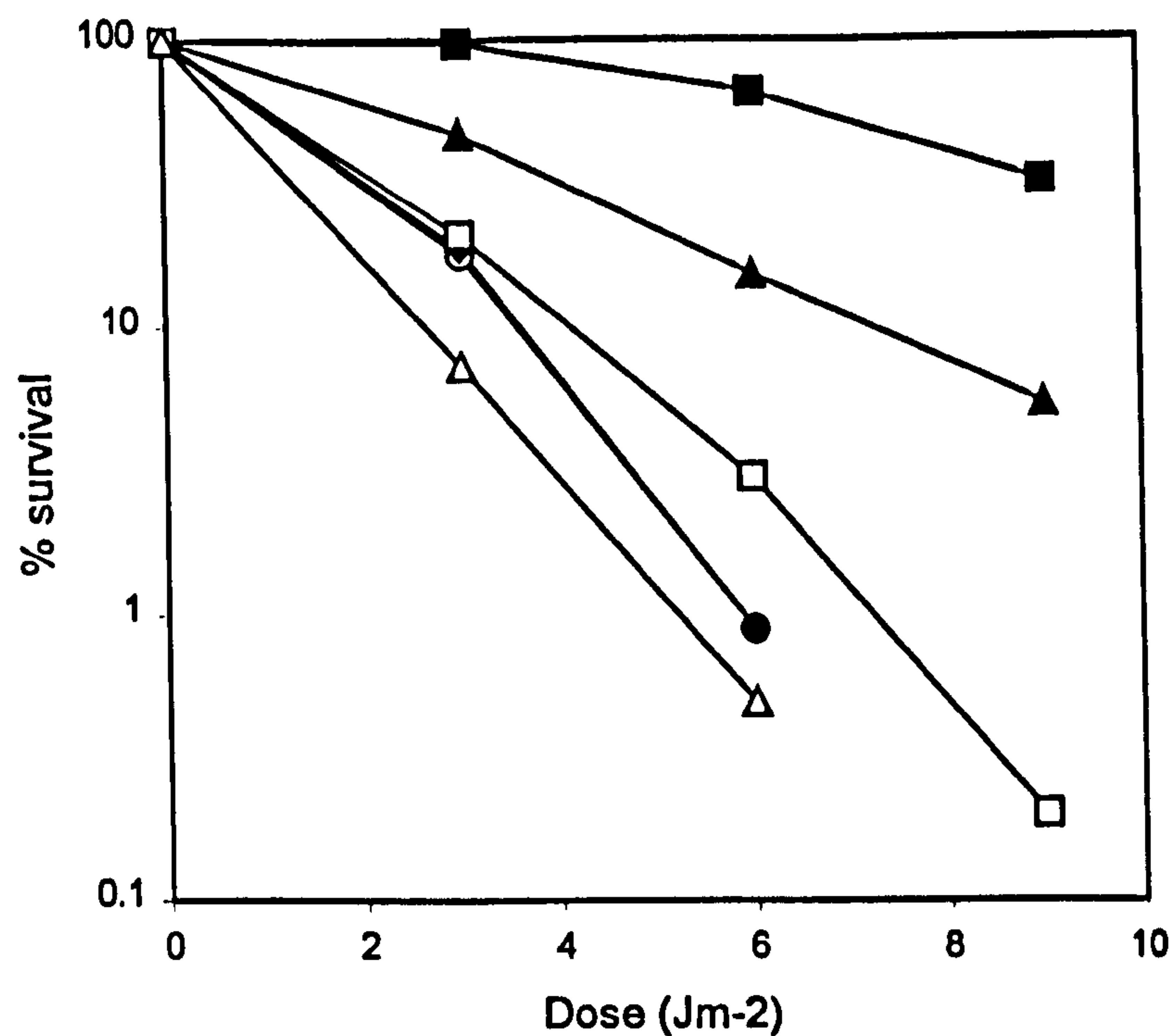


Figure 4.16 Graph illustrating the ability of *CSB* cDNAs with the mutations R670W (□), W851R (△) and V957G (◆) to correct the UV sensitivity of UV61. All three mutants fail to show any survival above that of untransfected UV61 (○). ■ = AA8, ▲ = WT corrected.

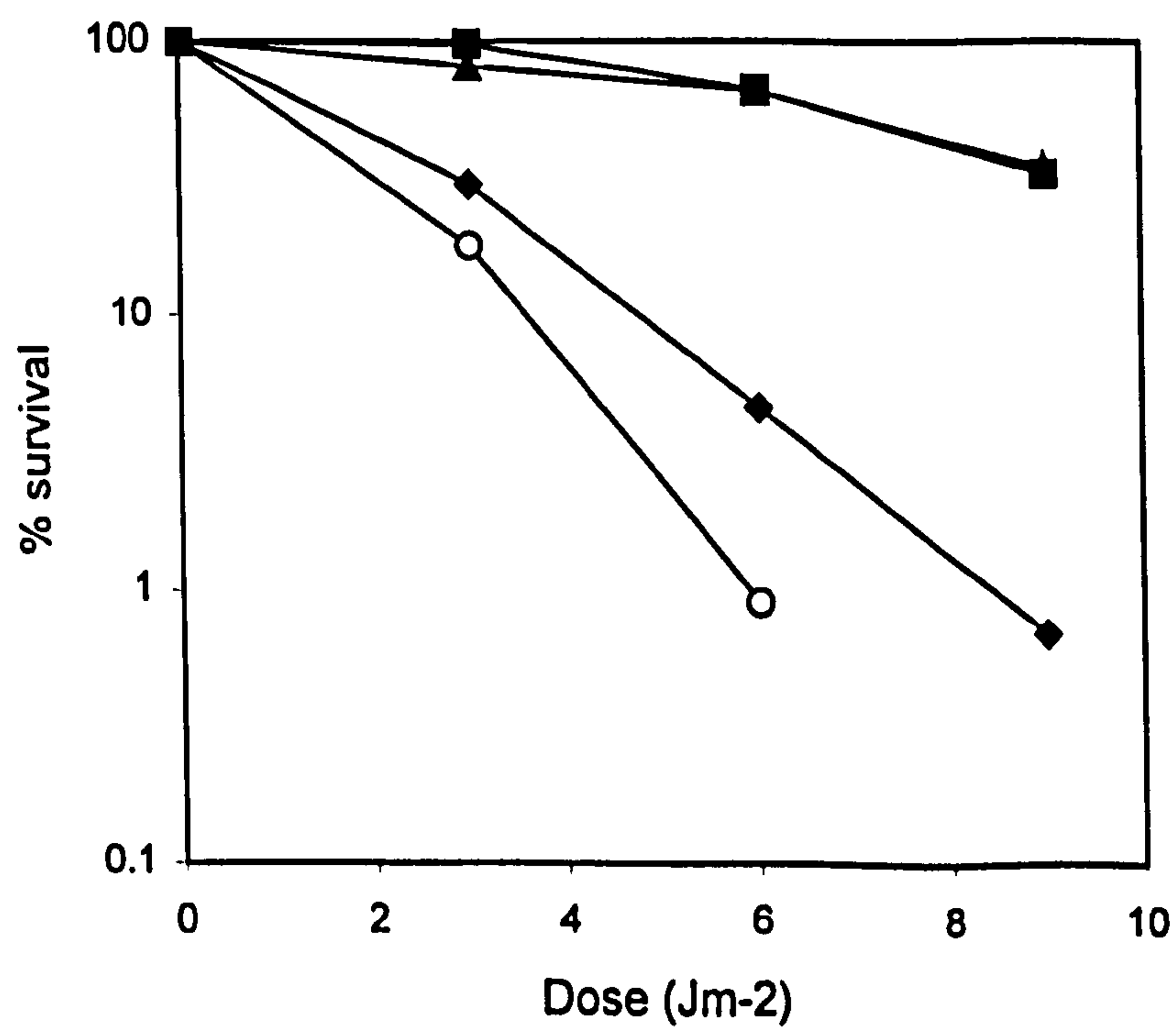


Figure 4.17 Graph illustrating the inability of *CSB* cDNA carrying the mutation P1042L (◆) to correct the UV sensitivity of UV61. The mutant fails to show any UV resistance above that of untransfected UV61 cells (○). ■ = AA8, ▲ = WT corrected.

levels up to AA8 parental levels. The aim of this study was to determine whether mutations inactivated the gene, therefore full correction by the wild-type gene was not essential for the purposes of this experiment. The mutants R670W, V957G and W851R were cloned into pCDNA3 for transfection and as illustrated in Figure 4.16, all fail to show any significant levels of UV resistance above that of the UV61 cells either untransfected or carrying empty vector.

In contrast to the partial complementation seen with *CSB* carried on pCDNA3, when the gene was transfected on the alternative expression vector pCI-neo, it was able to restore UV resistance back to AA8 parental levels (Figure 4.17). The last of the mutations tested for ability to correct, P1042L failed to show any correction of the UV sensitive phenotype of UV61 (Figure 4.17). Therefore, all amino acid substitutions identified in Cockayne Syndrome patients as putative causative mutations are unable to correct the UV sensitivity of the mutant UV61 hamster cells. It can be concluded that they each result in the production of protein in which the DNA repair function has been completely inactivated.

4.5 Discussion

The analysis of eleven patients from different backgrounds has revealed a wide variety of mutations in the *CSB* gene. A considerable number of the mutations found in CS-B patients resulted in severely truncated products. Several patients possessed two alleles affected in this way and it is unlikely that any functional protein is produced. A similar result was also found in the patient CS1AN analysed by Troelstra et al., 1992. This confirms that *CSB* is not an essential gene, as also demonstrated for the *S. cerevisiae* homologue *RAD26* (van Gool et al., 1994). The generation of a CS-B knockout mouse (van der Horst et al., 1997) gives further support to the non-essential nature of the gene with the mice demonstrating near normal development.

The results of the analysis carried out on the CS-B patients are summarised in Table 4.2.

Cell line	Inactivating Mutation ^a	aa change	Other mutations ^b									
			T213C L45P	G214C L45L	A834C K255T	G1275A G399D	C2904T T942M	C2830T G917G	G2905C T942T	A3222G H1048R	A3368G M1097V	A4317G Q1413R
CS2TAN (88) YK	G1630A	Trp517stop		Hom								
CS1TAN (77) YA	C2282T	Arg735stop						Hom				
25627 JM	C2282T	Arg735stop										
	C1436T	Arg453stop										
CS4BR (39) HE	C629T	Gln184stop				Het						
	C2087T	Arg670Trp										
CS2BI (9) MS	C2087T	Arg670Trp						Hom				
	C3204T	Pro1042Leu										
CS1BE	C2087T	Arg670Trp						Het				
	-A3615	-1 1180										
CS3TAN (98) MI	T2630A	Trp851Arg		Hom								
CS1IAF (67) IN	T2949G	Val957Gly				Hom						
CS1ABR (107) RC	-G2218	-1 714										
CS2BE	-T3610	-1 1178							Hom			
	Del Exon 10	Exon 10 del										
CS1MA (34) LW	Del Exon 10	Exon 10 del 665										
CS2GO (113B) CF	T2630A	Trp851Arg	Het								Hom	Hom
	A3716G	Arg1213Gly										
CS3GO (113A) LF	A3716G	Arg1213Gly									Hom	Hom
	Del 3858-4653	Del 1260-1493										

^aMis-sense mutations shown to be inactivating by failure to restore UV resistance to UV61 cells. ^bMutations shown or assumed not to be inactivating. Hom, homozygous, Het, heterozygous.

Table 4.2 The mutations and polymorphisms found in CS-B patients during this study.

Of the 14 inactivating mutations identified, five were CG to TA transitions at CpG sites, presumably as a result of deamination of 5-methylcytosine. The proportion of such transitions is consistent with that found in other human genetic disorders. The other mutations present were a further two transitions, two transversions, two occurrences of aberrant splicing, two frameshifts and an insertion of unknown origin. The frameshifts shown here all result from the deletion of a single nucleotide from within a run of several repeated bases. Such deletions are relatively common and generally occur as a result of replication slippage.

The mutations identified did not reveal any regions within the gene that could be termed as hotspots. There is, however a tendency for the mutations to be located towards the 3' two thirds of the gene, indicated by the slight clustering of the mutations in this region. It can therefore be tentatively postulated that mutations occurring in the N-terminal part of the gene may have less pronounced effects on function.

When studying the effects of mutation location on protein function, it is only relevant to look at mis-sense mutations. Nonsense mutations and frameshifts result in truncated protein and therefore location will only influence protein size. Many of the mis-sense mutations are located in or around the helicase/ATPase motifs (Figure 4.2) suggesting that these domains are functionally significant. It has been demonstrated that although CSB has no detectable helicase activity it does have a strong DNA-dependent ATPase activity (Selby and Sancar, 1997; van Gool et al., 1997). These ATPase motifs are present in all members of the SWI/SNF superfamily, with many possessing ATPase rather than helicase activity. In the case of CSB and other proteins in this family it is more appropriate to refer to these domains as ATPase consensus motifs.

When the patient analysis was first begun, it was hoped that the location and nature of the mutations would show some correlation with the clinical features of the patient. Clinically, Cockayne Syndrome has been classified into two groups by Lowry, 1982. Patients with

type I CS appear normal at birth with symptoms appearing during the first few years. In contrast, type II patients exhibit severe symptoms at birth and show little or no physical or neurological development. Of the patients studied, nine are type I and two are type II CS. When considering the relationship between the nature of mutations and the clinical features it is interesting to look at the four patients 25627, CS1TAN, CS1MA and CS1ABR. All four patients possess termination codons in both alleles of *CSB* and are unlikely to produce any functional protein. Two of these, CS1TAN and 25627 are type I CS and the others CS1MA and CS1ABR are type II. Therefore, carrying mutations resulting in non-functional CSB protein can have a wide variety of symptoms and be classified as either form of CS. The variation of clinical symptoms exhibited by CS patients must be a result of some other environmental or developmental stimulus.

The findings of this study have led to a number of hypotheses about important domains and regions of the *CSB* gene. The clustering of the mutations towards the 3' end and the high levels of conservation in the central part of the gene prompted a study into the functional significance of the N- and C-terminal ends of the protein (chapter 6). Also, the presence of a highly acidic region of amino acids and a stretch of glycine residues led us to study the effects of removing and replacing these regions (chapter 6).

With CSB showing significant homology to the yeast SNF2 gene and possessing strong ATPase activity, it was thought that the seven helicase/ATPase motifs may prove crucial to protein function, particularly with the high number of mutations occurring within this region. To this end, a number of mutations were created in this region analogous to mutations created in SNF2 (Peterson et.al.,1997) and will be tested for functional complementation in UV61 and for activity in *in vitro* assays (see chapter 7 for greater detail).

5.1 Introduction

During the study of the mutations in the *CSB* gene of Cockayne Syndrome patients, the siblings CS2GO and CS3GO proved to be a more interesting and complex case than the other patients studied. They are a brother and sister of Scandinavian origin and are the offspring of unaffected parents who are not known to be related.

Homozygous mutations result from the same affected allele being inherited from each parent, i.e. the alleles are of common ancestral origin. Therefore, when such mutations occur the parents of the affected patient are often related and the mutation has arisen in an earlier generation of the family to which they both belong. In contrast, compound heterozygous mutations are the result of inheriting a different affected allele from each parent. The two mutant alleles are likely to be from different origins arising as new mutations either in the patient or within earlier generations of each family. The parents of patients in these cases are often unrelated and have separate genealogy.

The siblings studied here carry compound heterozygous changes and it was anticipated that the pattern of mutations identified in both the brother and sister would be identical. Since the parents are unaffected, presumably carrying only one affected allele each, to produce affected offspring the child must inherit the one affected allele from each parent. However, the siblings presented here do not show the expected pattern of mutations.

5.2 Materials and methods

For the main part of this study, the same techniques of RNA extraction, RT-PCR and direct sequencing were employed as outlined in chapter 4 (section 4.2). In addition to the standard techniques a PCR and cloning analysis was carried out. To determine the allelic

origin of mutations occurring at the C-terminal end of the gene PCR fragments V and VI were amplified in one piece and cloned into T vector. The resulting clones were sequenced by standard Sequenase sequencing of ssDNA to determine presence of mutations in the different alleles. PCR product I was also cloned and sequenced using the T vector system to confirm the presence of a possible mutation.

5.3 Results

The first of the siblings to be sequenced was the brother CS2GO, yielding five amino acid changes. Initially, two mis-sense mutations were identified, the first of which had been seen previously in the patient CS3TAN as described in chapter 4. The mutation is a T to A at nucleotide 2630 and is shown in figure 5.1, seen as heterozygous in this patient, compared to the wild-type sequence, and that of the homozygous CS3TAN. This results in the substitution of tryptophan 851 with an arginine and is located at the edge of helicase domain IV within 37 aas showing 76% identity to Rad26.

The second mutation found in CS2GO was a C to T transition at nucleotide 213, the most N-terminal change seen to date in *CSB*. The mutation is heterozygous and was only just visible on the sequencing gel. To confirm this observation, PCR fragment I was cloned into T vector and subsequent clones were analysed for presence of the mutation. Both wild-type and mutant clones were observed and figure 5.2 shows the sequencing of one clone of each, confirming the presence and heterozygosity of C213T. This change results in a leucine to proline change at residue 45. Although the N-terminus of the protein is not particularly well conserved (appendix II), the introduction of a proline residue is likely to result in conformational changes in the protein.

Previously when two potentially inactivating mutations have been found in a CS-B patient, further sequence analysis has not been carried out. However, because of the unusual extreme N-terminal position of the L45P mutation, it was decided that sequencing of the

rest of the cDNA was necessary. This full sequence revealed that a full-length cDNA was obtained and that the sequence was identical to the wild-type sequence (except for the substitution that occur in the C-terminal third of the gene).

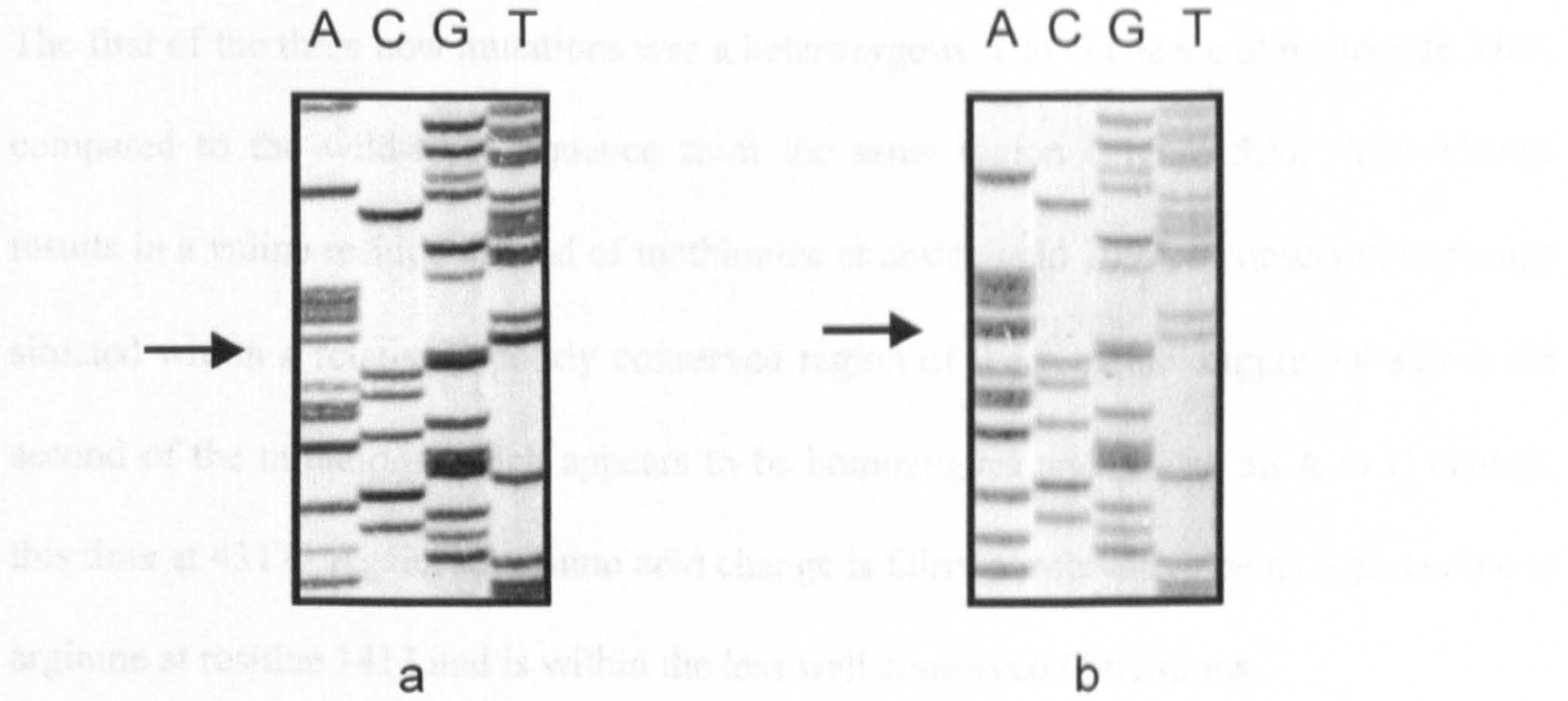


Figure 5.1 Sequence of wild-type (panel a) and patient CS2GO around position 2630 showing the presence of both the wild-type T and mutant A nucleotides at the same position in panel b.

(Figure 5.3). The A37163 change results in the substitution of an arginine residue for a glycine residue. Although the overall region is well conserved, the substitution of an arginine residue for a glycine residue is a significant change. The amino acid change is the opposite of the one observed in the mutant.

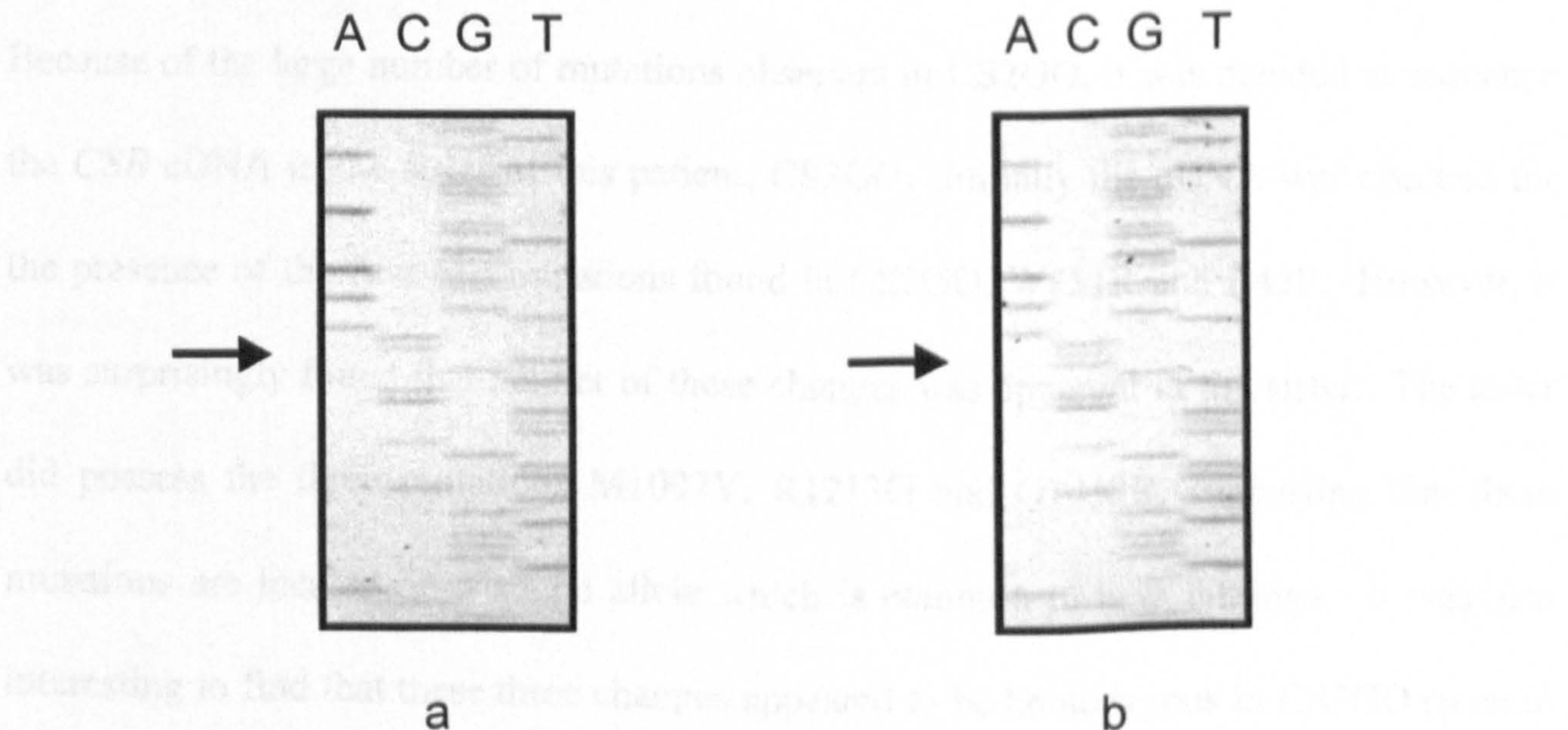


Figure 5.2 Sequence of a wild-type (panel b) and a mutant (panel a) clone around position 213 showing both alleles from patient CS2GO. The wild-type has a C and the mutant a T at nucleotide 213.

originates a nucleotide 385K and continues to 400K where it is located in the wild-type sequence. Figure 5.4 shows the sequence of the mutant clone around position 213.

rest of the cDNA was necessary. This full sequence revealed three additional amino acid substitutions that occur in the C-terminal third of the gene.

The first of the three new mutations was a heterozygous A to G change at nucleotide 3368, compared to the wild-type sequence from the same region (Figure 5.3). This change results in a valine residue instead of methionine at amino acid 1097, a conservative change situated within a relatively poorly conserved region of the protein. Figure 5.4 shows the second of the mutations which appears to be homozygous and is also an A to G change, this time at 4317. Again, the amino acid change is fairly conservative being a glutamine to arginine at residue 1413 and is within the less well conserved C-terminus.

The final mutation observed in CS2GO is an A to G at nucleotide 3716 and can be seen clearly as heterozygous on sequencing of this region when compared to the wild-type (Figure 5.5). The A3716G change results in the substitution of arginine 1213 with a glycine residue. Although the overall region is not exceptionally well conserved, the precise residue is retained between human and mouse, although not present in yeast, and the amino acid change is the most severe of the three observed.

Because of the large number of mutations observed in CS2GO, it was decided to sequence the *CSB* cDNA in the sister of this patient, CS3GO. Initially the cDNA was checked for the presence of the first two mutations found in CS2GO, W851R and L45P. However, it was surprisingly found that neither of these changes was apparent in the sister. The sister did possess the three mutations M1097V, R1213G and Q1413R, suggesting that these mutations are located on a single allele which is common to both siblings. It was also interesting to find that these three changes appeared to be homozygous in CS3GO (seen in Figures 5.3, 5.4 and 5.5).

The mutation in the second allele of CS3GO is a C-terminal deletion. The deletion originates at nucleotide 3858 and stretches to 4563, which is beyond the end of the open reading frame. Figure 5.6 shows the sequence obtained from direct sequencing of fragment

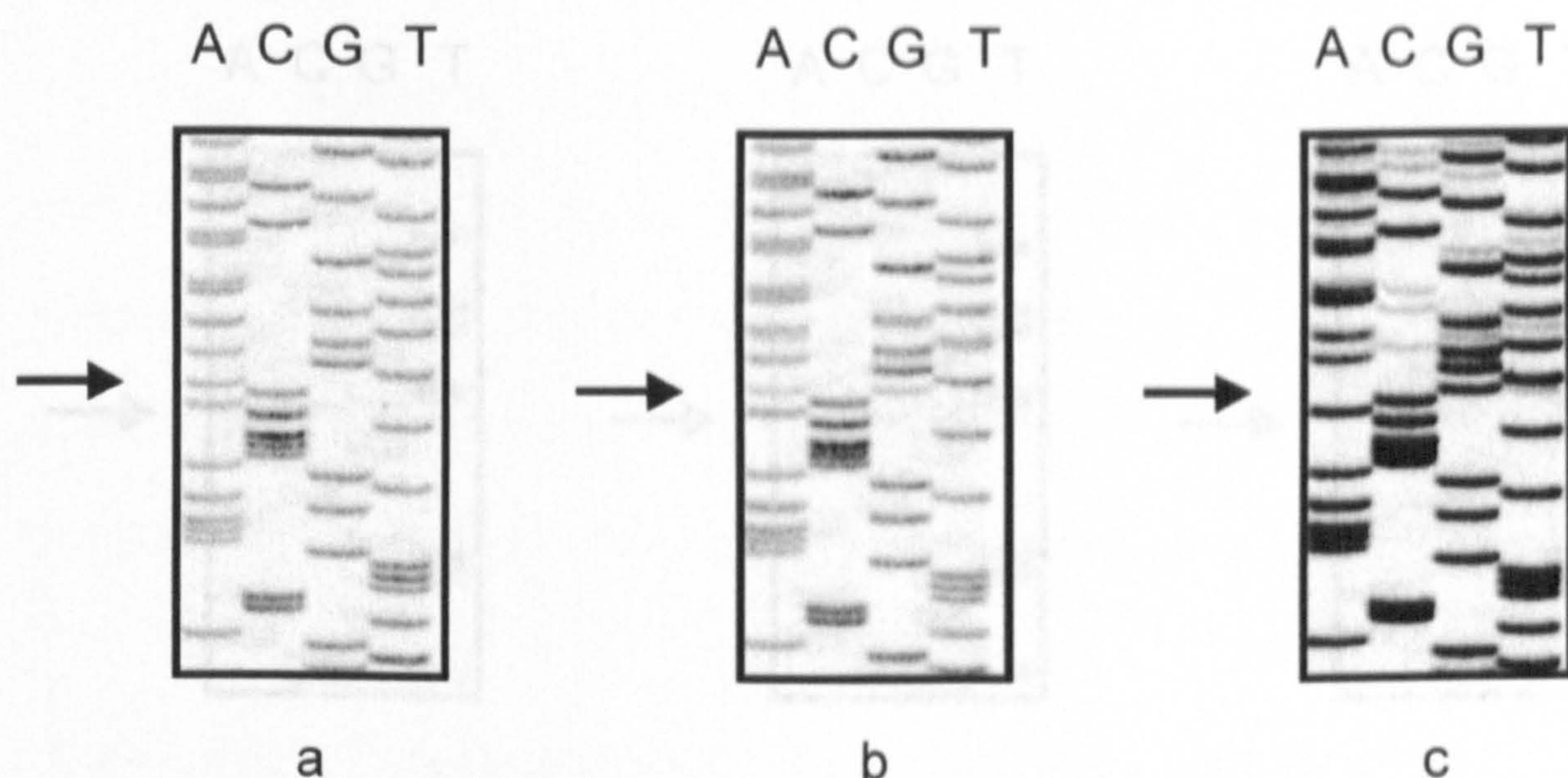


Figure 5.3 Sequence of wild-type (panel a) and patients CS2GO (panel b) and CS3GO (panel c) around position 3368. CS3GO possesses a mutant G nucleotide in the place of the wild-type A, whilst panel b shows the presence of both mutant and wild-type alleles in CS2GO.

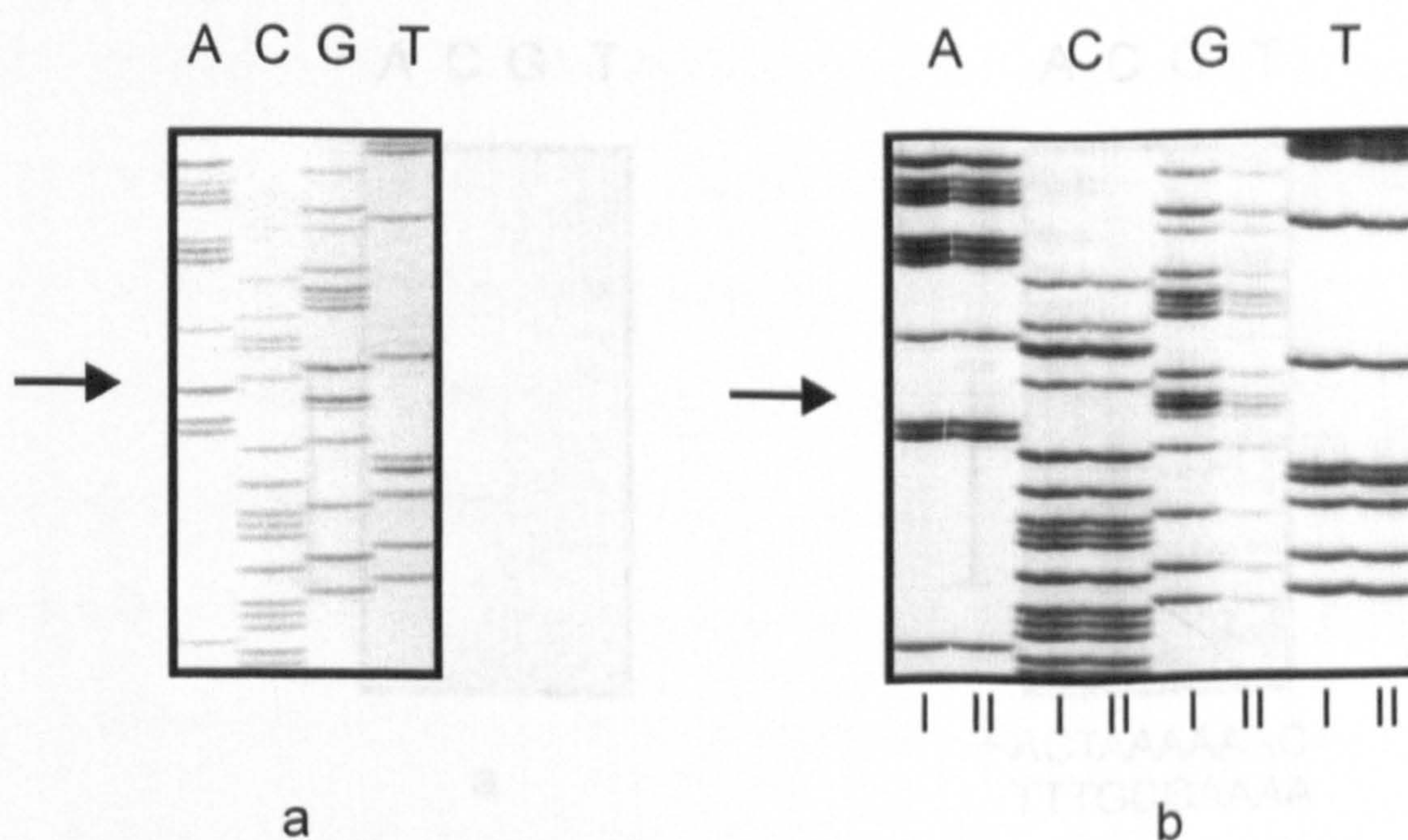


Figure 5.4 Sequence of wild-type (panel a) and mutant (panel b) around position 4317. Panel b shows patients CS2GO (lanes I) and CS3GO (lanes II) who possess a G in the place of the wild-type A nucleotide at 4317.

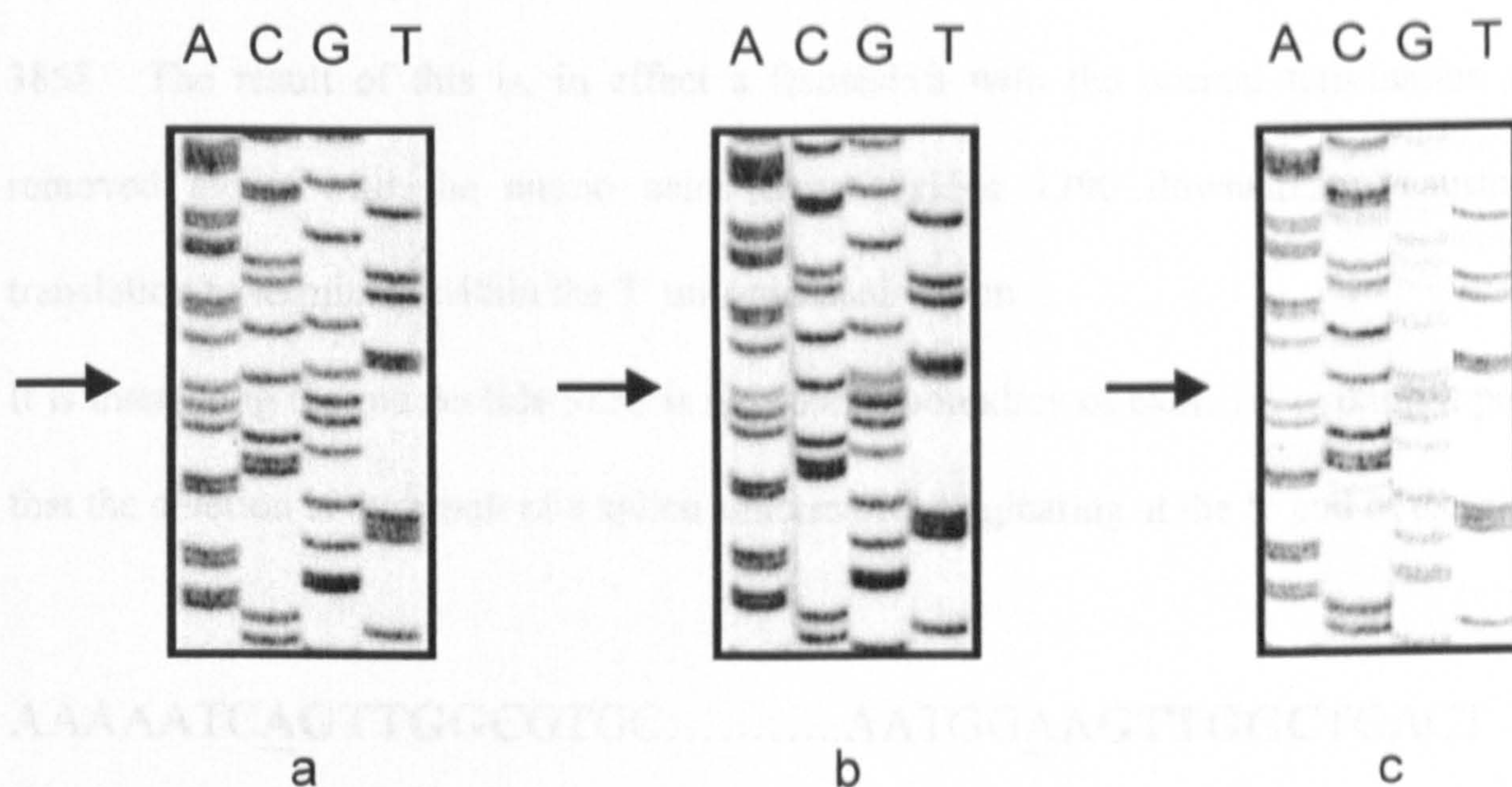


Figure 5.5 Sequence of wild-type (panel a) and patients CS2GO (panel b) and CS3GO (panel c) around position 3716. Panel c shows CS3GO to have a mutant G nucleotide in place of the wild-type A, whilst panel b shows the presence of both the wild-type and mutant alleles in CS2GO.

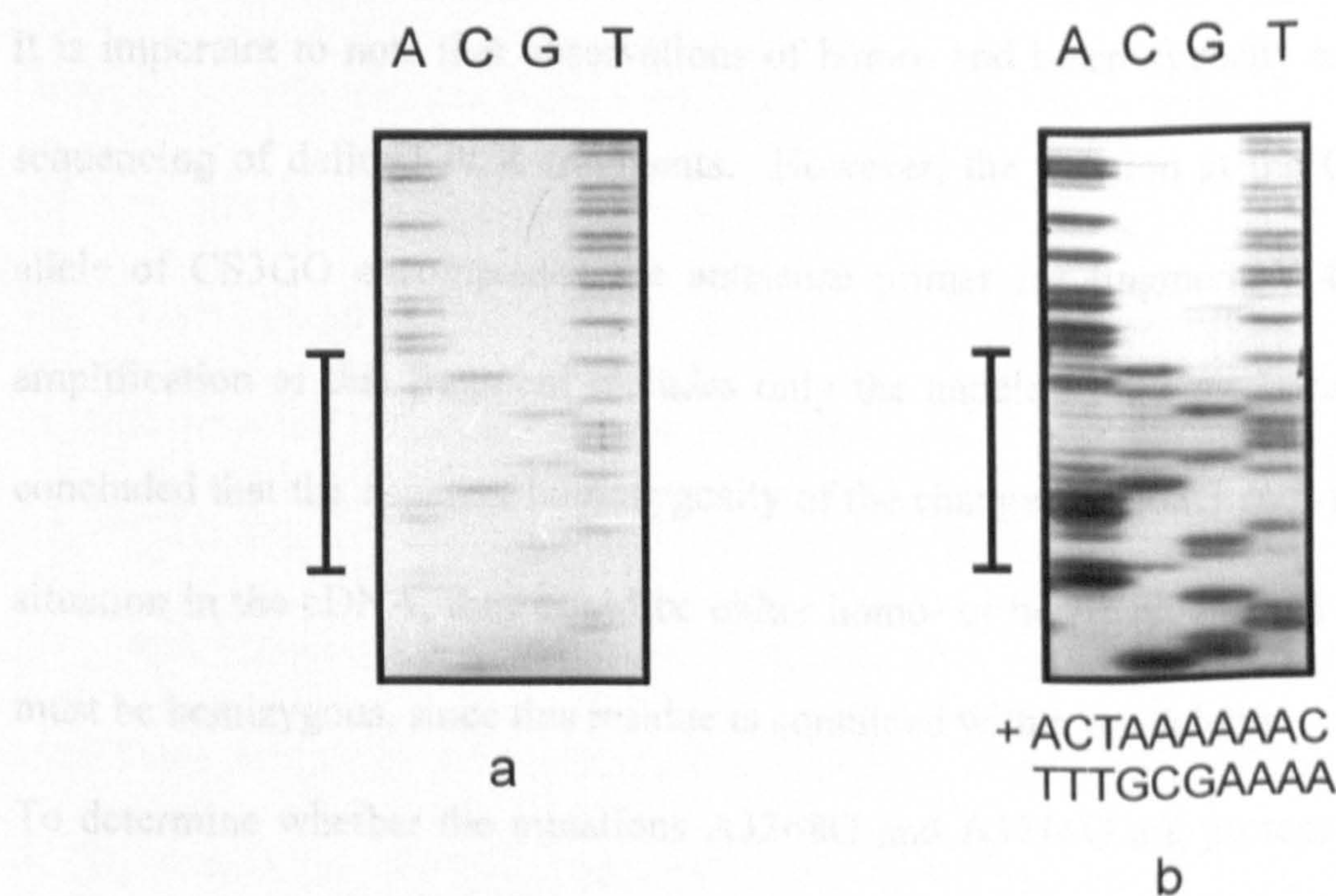


Figure 5.6 Sequence of wild-type (panel a) and patient CS3GO (panel b) around position 4561. The marked region in panel b shows the presence of both the wild-type and deleted alleles. The mutant sequence is that of nucleotide 3858 to the 5' end of PCR fragment six(primer VIB-bio) reading 3' to 5'.

VI using the reverse primer W4679 located at nucleotides 4696-4679. The start of the sequence is normal, then at 4653 an alternative sequence appears, which reads 3' to 5' from 3858. The result of this is, in effect a frameshift with the normal termination codon removed along with the amino acids from residue 1290 downstream, causing the translation to terminate within the 3' un-translated region.

It is interesting that nucleotide 3858 is also the 5' boundary of exon 18, making it possible that the deletion is the result of a splice abnormality originating at the 5' end of exon 18:

AAAAATC**A**GTTGGCGTGC.....AATGGA**A**AGTTGGCTGACT

Underlined are the boundaries, 3858 and 4653 of the deletion and bold indicates the repeated sequence.

It is however more likely to be due to polymerase slippage with a repeat of six nucleotides occurring at each end of the deletion. Deleting such a large portion of *CSB* is very likely to result in a non-functional protein.

It is important to note that observations of homo- and heterozygosity are based on direct sequencing of defined PCR fragments. However, the deletion at the C-terminus of one allele of CS3GO encompasses the antisense primer for fragment V(4253 to 4230), so amplification of this fragment includes only the undeleted allele. It cannot therefore be concluded that the apparent homozygosity of the changes A3368G and A3716G reflect the situation in the cDNA, they could be either homo- or heterozygous. In contrast, A4317G must be hemizygous, since this residue is contained within the deletion.

To determine whether the mutations A3368G and A3716G are present in both alleles of each patient, the cDNA was amplified using the sense primer for fragment V and the antisense primer for fragment VI. This yields a fragment of approximately two kilobases, which was cloned into a T vector and subsequent clones sequenced to determine the

presence or absence of the two mutations. The sequencing revealed clones possessing either both the wild-type or both mutant sequences demonstrating that the second allele is wild-type at these positions, and that these mutations are in fact heterozygous.

5.4 Functional Analysis

With the large number of mutations present in these two patients it was important to determine which were the inactivating mutations in each allele. The alleles unique to each patient carry mutations that are very likely inactivating, the deletion in CS3GO and W851R in CS2GO, shown to be inactivating for CS3TAN in chapter 4. However in the case of the first allele of CS2GO one problem remains, the presence of L45P, which might be expected to result in inactive protein. To determine the effect of L45P and to discover the inactivating mutation in the common allele it was necessary to perform a functional analysis of these changes.

The two mutations in allele one of CS2GO were created in pCDNA3 and transfected into UV61 as described in chapter 2. As noted before the wild-type gene carried on this vector is unable to fully correct the UV sensitivity of the hamster cells. Figure 5.7 shows the survival curves generated for wild-type, L45P and W851R cDNAs in clones of UV61. This clearly shows that L45P has no effect on the DNA repair function of CSB since the survival is similar to that of cells carrying wild-type cDNA. In contrast as seen previously, the W851R mutation destroys the ability of the cDNA to correct.

The three mutations carried on the common allele were each create in pCI-neo and transfected into UV61 cells along with a wild-type. Clones generated from the transfection of cDNA with the Q1413R change produced the survival curve shown in figure 5.8. The wild-type transfectants correct the UV sensitive defect fully, as do the transfectants carrying cDNAs with the mutation Q1413R. The other two mutations demonstrate a far less clear cut result with neither cDNA able to correct the UV sensitivity of UV61 fully.

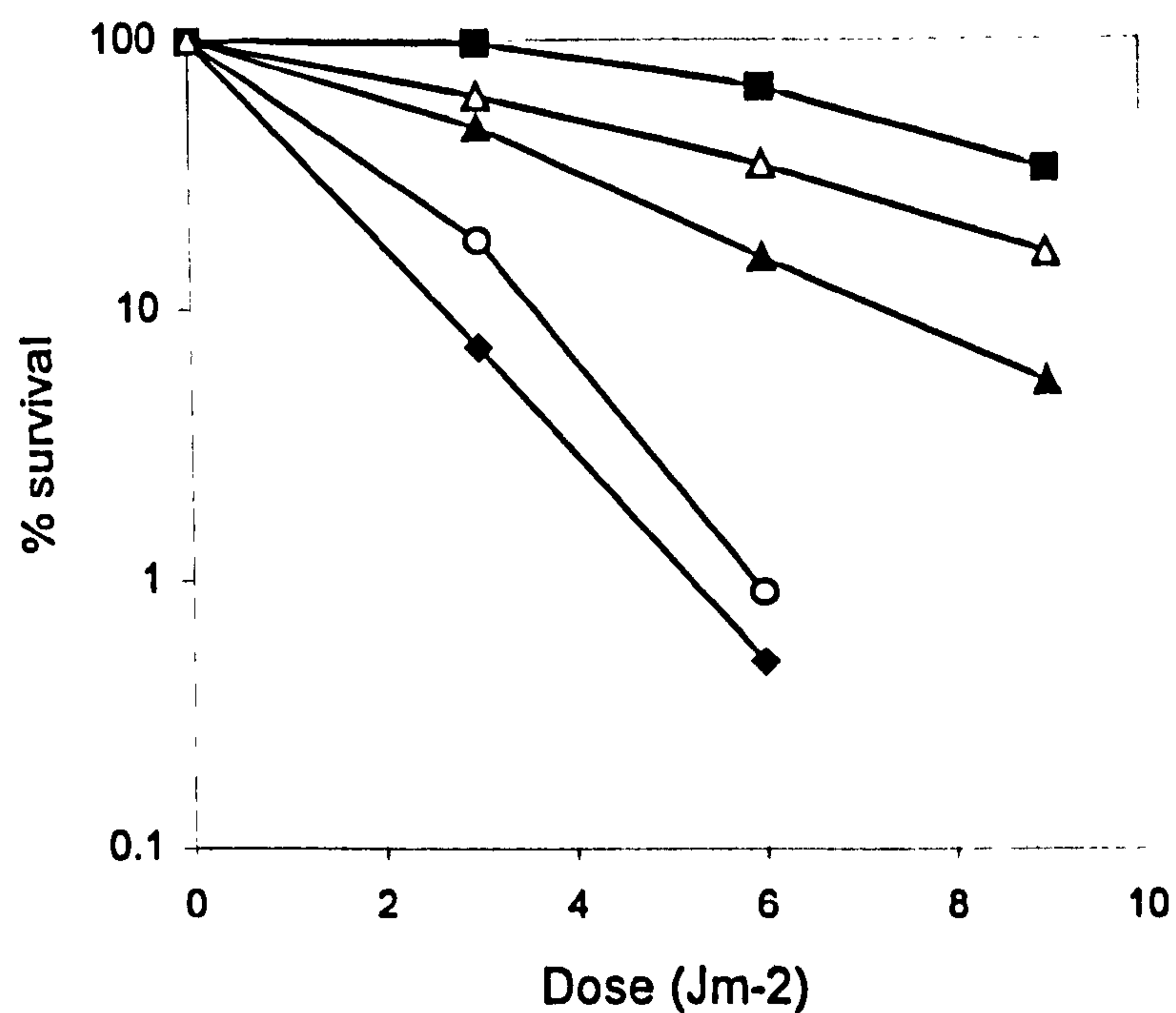


Figure 5.7 Graph illustrating the ability of *CSB* cDNAs with the mutations W851R (◆) or L45P (△) to correct the UV sensitivity of UV61. W851R fails to show any survival above that of untransfected UV61 cells (○), but L45P restore survival to the same level as wild-type cDNA (▲). ■ = AA8.

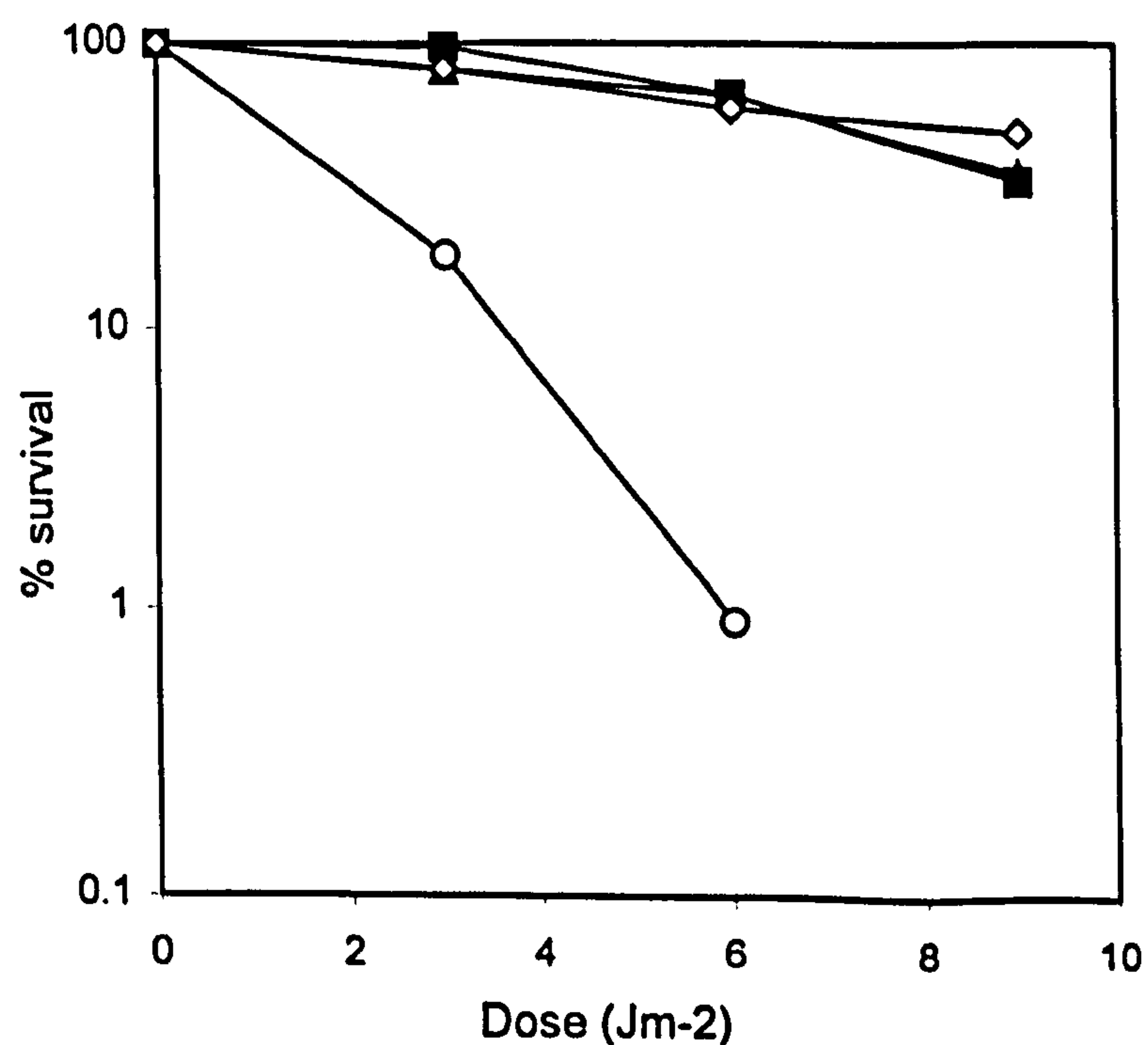


Figure 5.8 Graph illustrating the ability of *CSB* cDNA carrying the mutation Q1413R (◇) to correct the UV sensitivity of UV61 (○). The cDNA is able to restore UV resistance to the same levels as WT cDNA (▲), back to that of the parental AA8 cells (■).

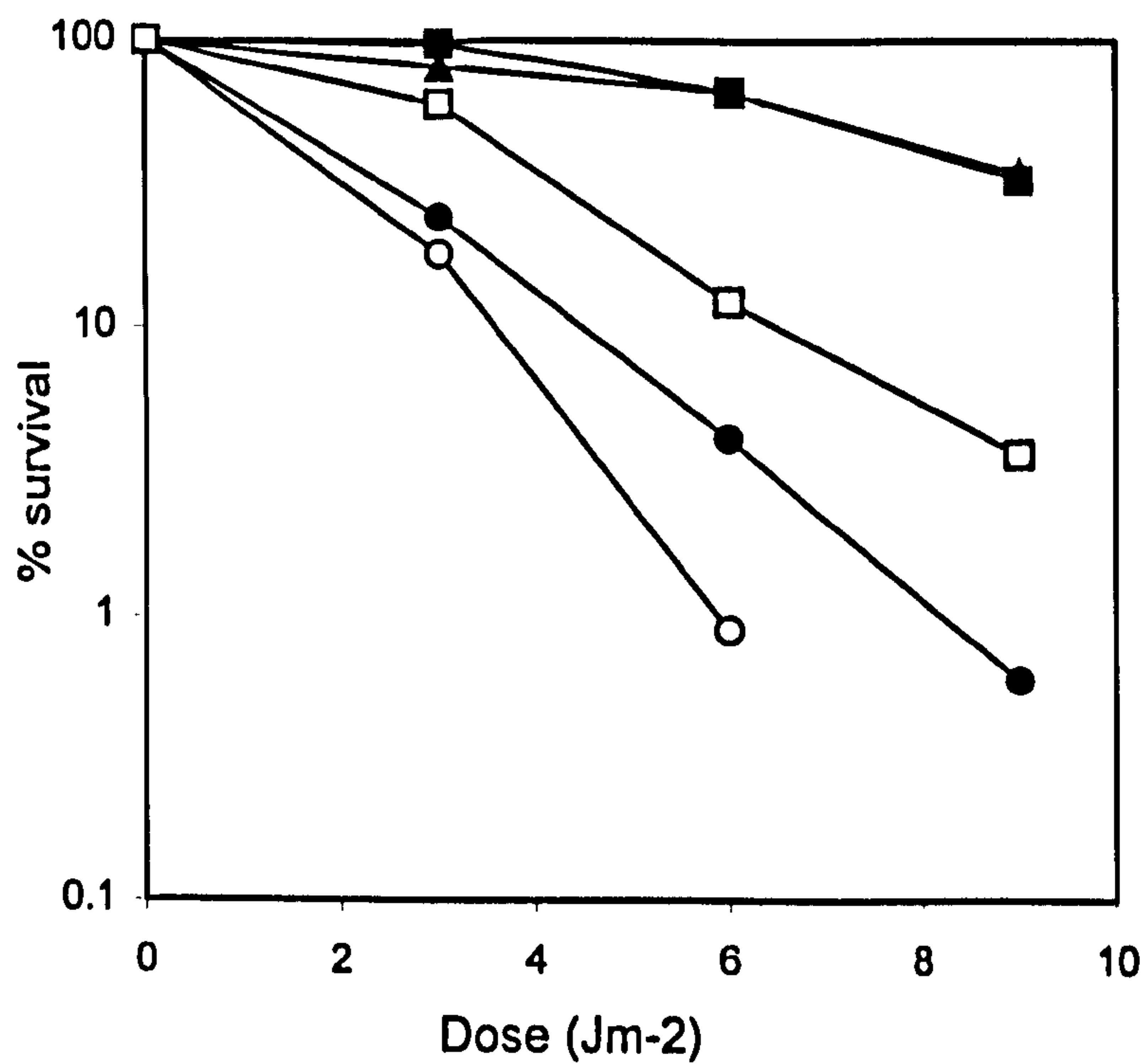


Figure 5.9 Graph illustrating the ability of *CSB* cDNA with the mutation R1213G to correct the UV sensitivity of UV61 (○). The mutant clones vary in their ability to correct, but both fail to correct to the same extent as WT cDNA (▲). ■ = AA8, ● = clone 2, □ = clone 10.

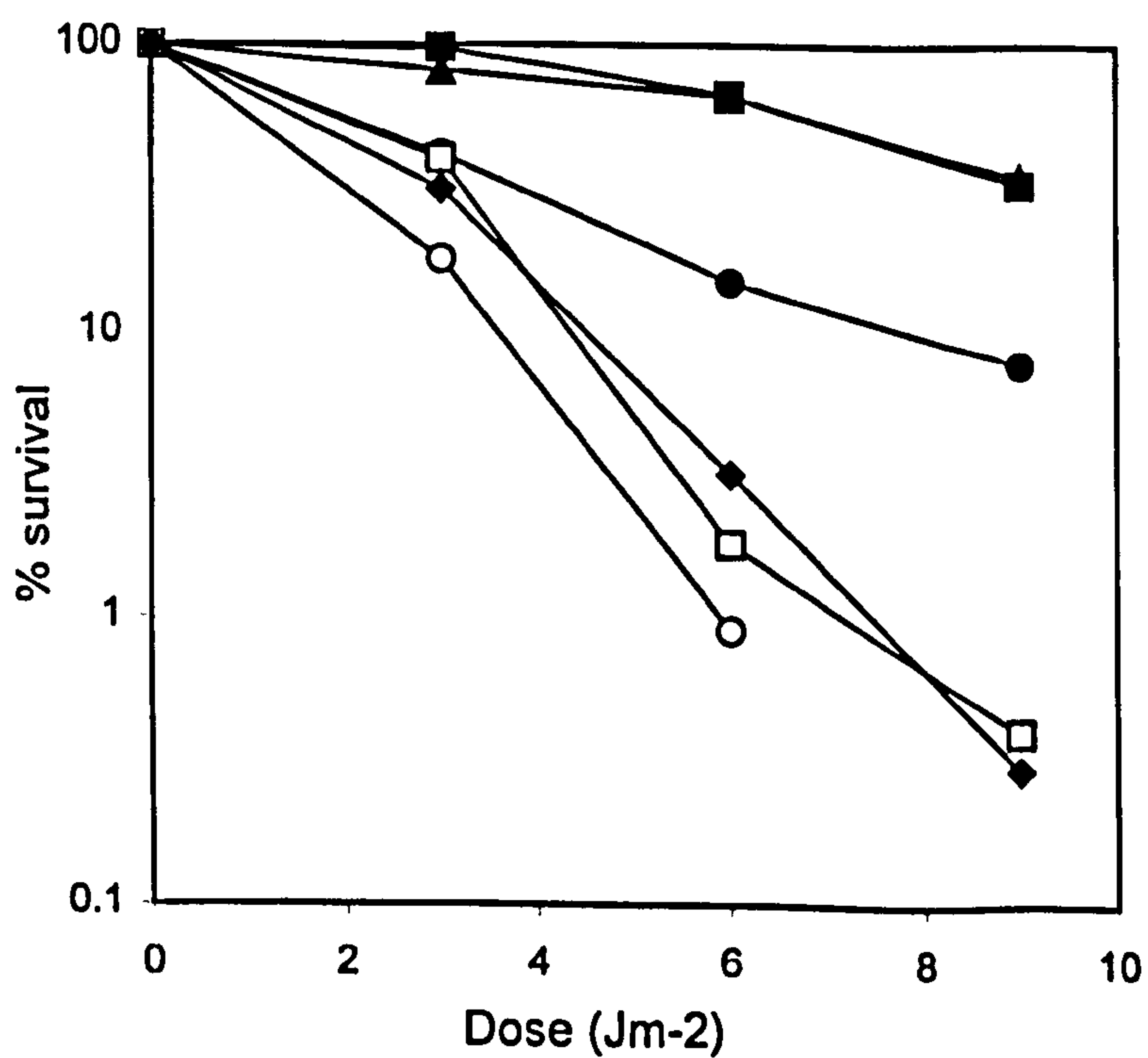


Figure 5.10 Graph illustrating the ability of *CSB* cDNA with the mutation M1097V to correct the UV sensitivity of UV61 (○). The mutant clones vary in their ability to correct, but all fail to correct to the same extent as WT cDNA (▲). ■ = AA8, ● = clone 1, □ = clone 4, ◆ = clone 5.

Two clones were analysed that carried the R1213G mutation, of which one showed no correction above UV61 levels, and the other was intermediate (Figure 5.9). Three clones carrying the mutation M1097V were also analysed for survival and both of these gave intermediate levels of correction (Figure 5.10). It is difficult to determine whether either or both mutations are responsible for the CSB phenotype of the patients CS2GO and CS3GO.

5.5 Discussion

The siblings CS2GO and CS3GO have presented an interesting case, both through the spectrum of mutations observed and the presence of three different mutant alleles. The brother and sister possess one common allele, presumably inherited from the one parent. Therefore the other two alleles should be inherited from the second parent, who theoretically has mutations in both alleles. However, neither parent is known to be affected with CS, so it must be concluded that either one of the siblings inherited the normal allele from one parent and a new mutation that has occurred in that child. Alternatively, the two siblings might have different biological fathers. The most conclusive way to determine this would be to screen the parents for the mutations observed in the children, to determine their origin. A request for material from the parents has been made, but currently none is available.

The patients CS2GO and CS3GO possess several different mutations, one deletion and five mis-sense mutations (shown in Figure 5.11). The mutations in the common allele M1097V, R1213G and Q1413R were tested for functional complementation in hamster cells to determine which of the three is responsible for the CS phenotype. The data obtained from the complementation analysis is inconclusive as to which mutation is phenotypically significant. It is clear that the Q1413R change does not affect protein function, as the cDNA with this mutation is able to restore the UV resistance of UV61 cells

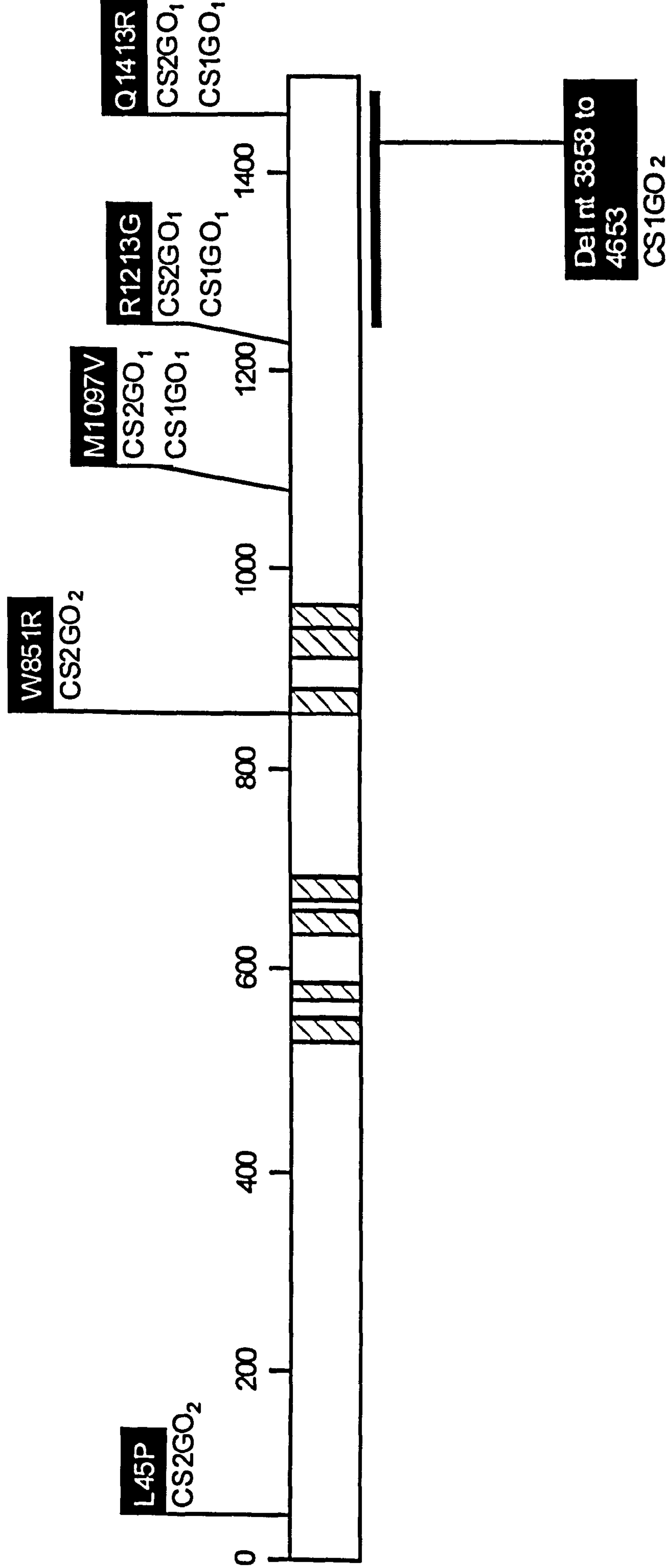


Figure 5.11 Diagrammatic representation of CSB illustrating the position and nature of mutations found in the CS-B patients CS2GO and CS3GO. The mutations are shown in the black boxes with the patient codes below, along with the alleles as subscripts. ▨ = helicase domains.

back up to wild-type levels. However, neither of the other two mutations is clearly inactivating or able to correct.

These three mutations occur together in two other patients, CS8BR and CS7TAN (analysed by our colleagues in Pavia), unrelated to each other and the patients studied here (Mallery et al., 1998). It is therefore possible that these mutations comprise a single haplotype or act additively. Alternatively it is possible that either mutation alone is inactivating and that the intermediate levels observed are a result of over-expression. Since the levels of expression of CSB in the transfected cells are likely to be higher than would occur *in vivo*, it is reasonable to suggest that the cDNAs might not correct if expressed at endogenous levels.

Further analysis of these mutations is required to determine which of the above theories is most likely. Firstly, more clones from each transfection need to be analysed to check whether the intermediate survivals are reproducible. Western blots on cell extracts from these clones should then be used to determine the levels of CSB present and whether there is a correlation between the survival and protein levels. Secondly, both mutations need to be introduced into the same construct and transfected into cells. Clones derived from this transfection would then show whether the levels of survival from clones with both mutations differs from that of each mutation alone.

The mutations demonstrated here have provided some insight into the functionally significant regions of the gene. During the earlier stages of the sequencing study it had been noted that mis-sense mutations clustered around the helicase/ATPase domains in the middle of the gene. It was postulated at this stage that the N- and C-terminal ends of the gene could be deleted without significantly affecting the protein function. The deletion seen in CS3GO which removes the C-terminal 233 amino acids is the only mutation in this allele and is therefore inactivating. This shows that removal of the C-terminus of CSB results in non-functional protein. The deletion seen here is quite large and would be expected to affect the protein, so smaller deletions are analysed for their effect on DNA

repair function in chapter 6. The next chapter also investigates the effects on N-terminal deletions on the DNA repair function of CSB. The change L45P gives us some insight as to what these effects might be, through its ability to correct the UV sensitivity of UV61. The introduction of a proline residue in the place of leucine might be expected to result in conformational changes in the protein because of the structural kink resulting from the introduction of this amino acid. This suggests that conformational changes in the N-terminus of CSB might not be significant for the DNA repair function, and that the region may not be essential.

CHAPTER 6 DELETION ANALYSIS

6.1 Introduction

Analysis of mutations in the *CSB* cDNA of patients with Cockayne Syndrome has provided information about the gene and important regions. Although a few sites were mutated in unrelated individuals and may therefore have arisen independently, there are no dramatic hotspots for mutation. In an attempt to discover further structure-function information, a series of deletions were created to determine whether certain regions of the gene were essential for function.

The majority of inactivating mis-sense mutations were clustered in the central two thirds of the gene, which is also the location of the seven helicase/ATPase domains. It was thought that the N- and C- terminal ends of the gene may not be required for the production of active protein. Information has been gained from the sequence analysis study through the L45P mutation and the 3858-4653 deletion. The L45P change does not have any effect on the DNA repair function of CSB, which is surprising as proline residues often result in conformational changes. However the deletion that removes the C-terminal end of the cDNA downstream of the 5' junction of exon 18 results in inactive protein. In order to investigate the involvement of the termini in repair function, two deletions were created each at the C- and N-terminal ends of the gene.

Two interesting domains were identified close to the N-terminus of the protein (Troelstra et al., 1992). Firstly, a stretch of six glycine residues and a glutamic acid residue at amino acids 440 to 446. Glycine residues are the smallest and uncharged amino acid residues and can often be spacer regions between highly charged or polar groups of residues. Secondly, at amino acids 356 to 395 is a region rich in acidic residues, which in particular contains seven consecutive glutamates (amino acids 378-384). Substantial concentration of negative charges might be expected to be important for protein function. To test

whether these domains really are important, they have been removed as part of this study and the deleted cDNAs checked for their ability to correct the UV sensitive phenotype of UV61 hamster cells.

6.2 Construction of deleted cDNAs

Six deletions were created in the *CSB* cDNA to carry out a structure-function analysis, and positions and lengths of these deletion are illustrated in figure 6.1. The sequences of all the primers used to create these constructs by PCR and site-directed mutagenesis are shown in Appendix I.

6.2.1 C-terminal deletions

Two C-terminal deletions were created removing 111 and 26 amino acids respectively. The smaller of these was created by site-directed mutagenesis of nucleotide 4481, replacing the wild-type G with an A. Therefore, a stop codon is introduced at amino acid 1468 in the position of the wild-type glutamic acid residue, resulting in the truncated cDNA I shown in figure 6.1, in which the C-terminal 26 amino acids are missing.

The larger C-terminal deletion was created by a restriction digestion method to give the deleted cDNA II in figure 6.1. The cDNA cloned into pcDNA3 was digested with the restriction enzyme *ApaI* which cuts at nucleotide 4227 in *CSB* and in the multiple cloning site of pcDNA3, 3' of the insert. This removes the C-terminal 111 amino acids including the stop codon and the 3' untranslated region. Termination therefore relies on the presence of a stop codon within the vector sequence, which occurs very close to the end of the *CSB* cDNA.

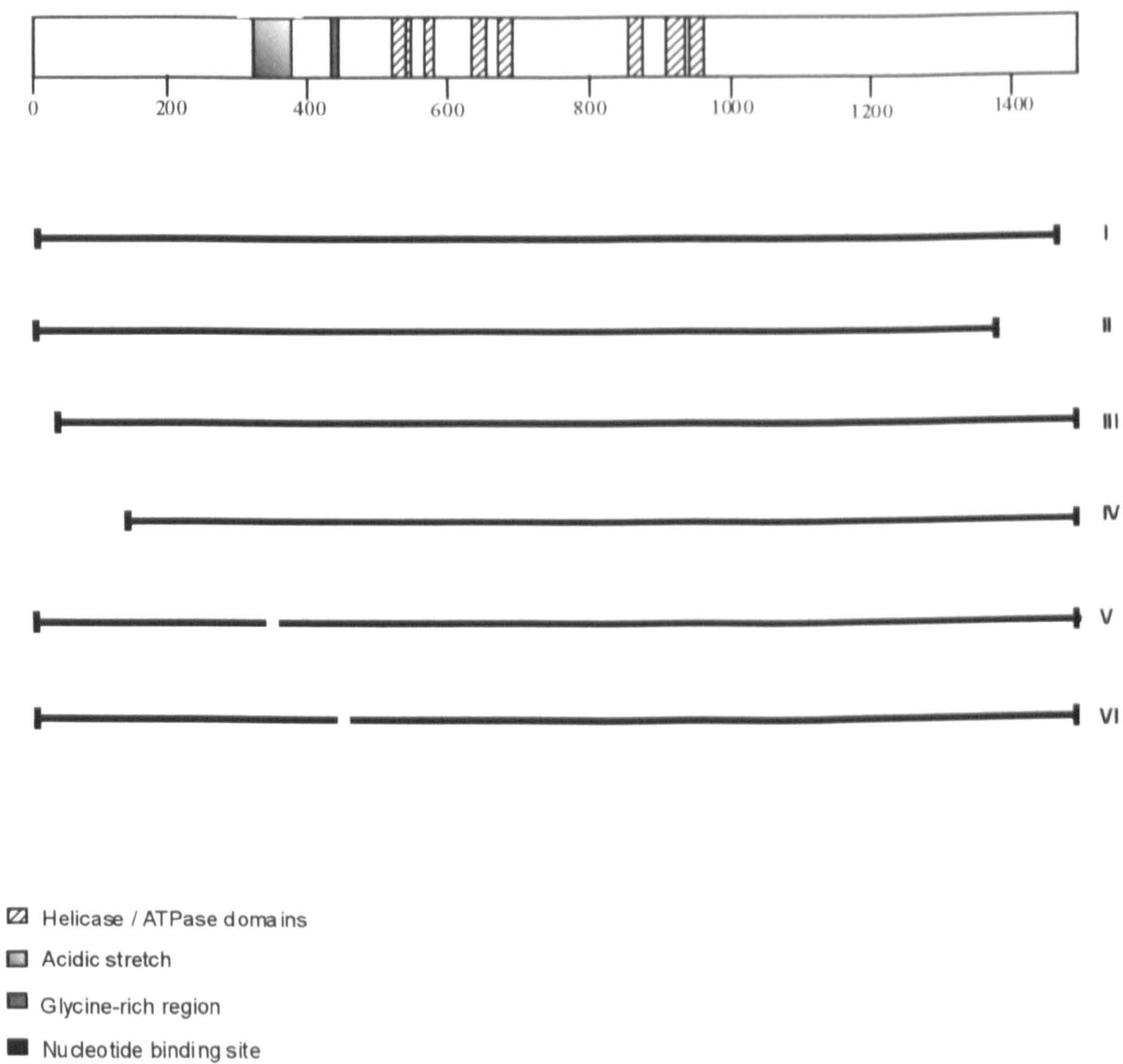


Figure 6.1 Diagrammatic representation of CSB showing the putative functional domains. The lines shown below the gene indicate the deleted cDNAs created and tested for functional complementation.

6.2.2 N-terminal deletions

Two N-terminal deletions were constructed using a PCR and cloning based method. PCR primers were designed to amplify the N-terminus of the cDNA, to create the deleted constructs III and IV of the lengths illustrated in figure 6.1, missing the first 28 and 137 amino acids respectively. The forward primer for amplification of the larger clone III was designed to cover the methionine residue at amino acid 29, using this as a transcriptional start site. To create the larger deletion, the forward PCR primer was designed to encompass a base change corresponding to nucleotide 491 which changes the leucine residue at 137 to a methionine, and therefore a start codon.

The resulting PCR products were cloned into the Invitrogen TA vector and sequenced to check for PCR errors. Error-free clones were digested with *SphI* at nucleotide 795 and *NotI*, in the polylinker 5' of the insert and cloned into *SphI/NotI* digested pcBlsSE6 (Figure 6.2). This replaces the full length wild-type N-terminal fragment with the newly created deletions. The new *CSB* cDNA constructs were transferred as *NotI/XhoI* fragments into *NotI/SalI* digested pCI-neo (Figure 6.2).

6.2.3 Internal deletions

Two deletions were created, each removing seven amino acids from the N-terminal third of the protein, disrupting either the glycine domain or the acidic rich region. Both deletions were created by site-directed mutagenesis of the full length cDNA in the mammalian expression vector pCI-neo. The sequences of the primers used for the mutagenesis are shown in Appendix I. The first deletion is located within a region of 39 acidic residues removing 21 nucleotides from 1211 to 1231 and consequently deleting seven glutamic acid residues at 378 to 384. The amino acid sequence of the acidic domain is shown below in Figure 6.3.

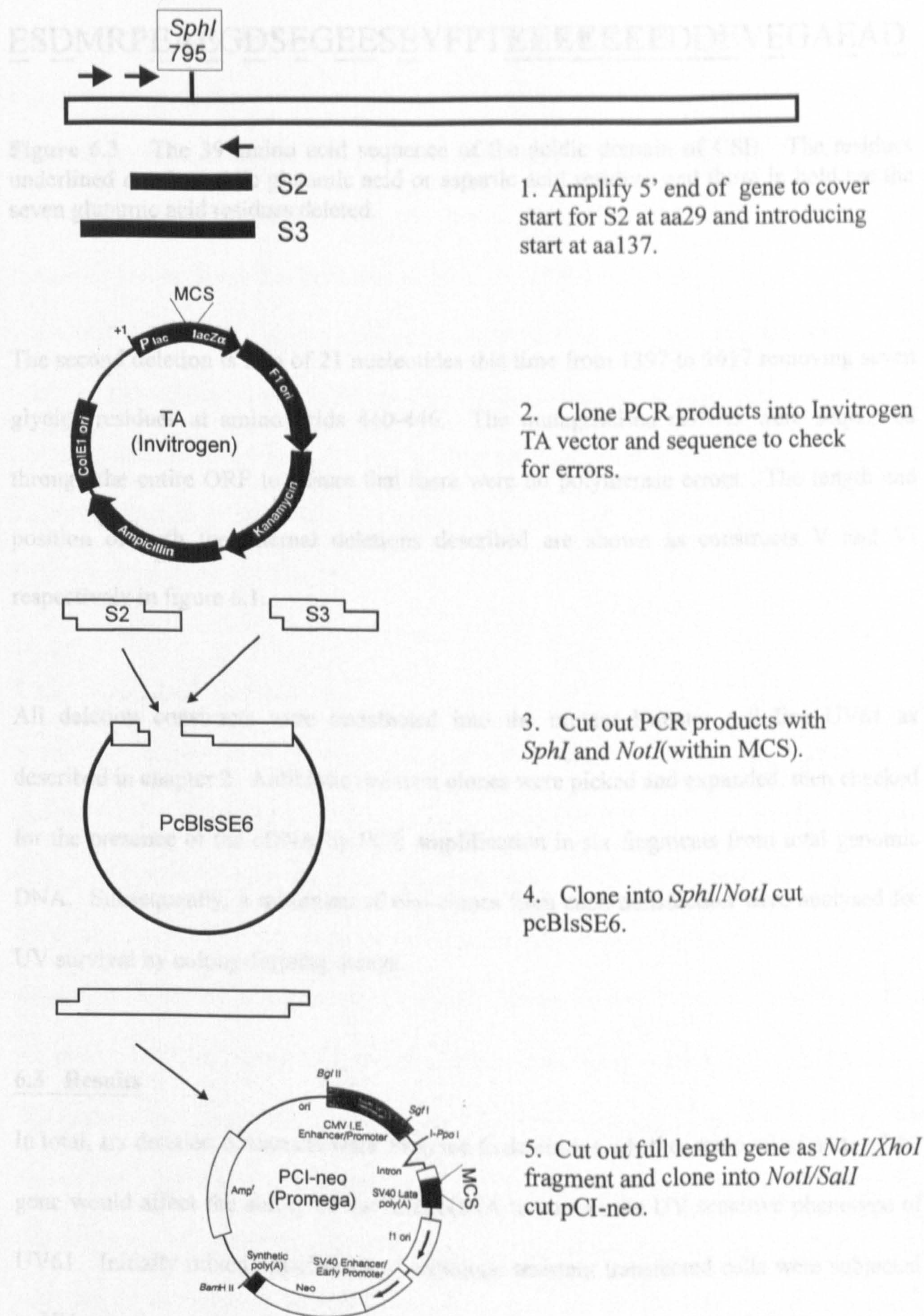


Figure 6.2 The cloning strategy for the creation of N-terminal deletions of CSB cDNA.

ESDMRPEAEGDSEGEESEYFPTEEEEEEEEEEDDEVEGAEAD

Figure 6.3 The 39 amino acid sequence of the acidic domain of CSB. The residues underlined are the acidic glutamic acid or aspartic acid residues and those in bold are the seven glutamic acid residues deleted.

The second deletion is also of 21 nucleotides this time from 1397 to 1417 removing seven glycine residues at amino acids 440-446. The mutagenised cDNAs were sequenced through the entire ORF to ensure that there were no polymerase errors. The length and position of both the internal deletions described are shown as constructs V and VI respectively in figure 6.1.

All deletion constructs were transfected into the mutant hamster cell line UV61 as described in chapter 2. Antibiotic resistant clones were picked and expanded, then checked for the presence of the cDNA by PCR amplification in six fragments from total genomic DNA. Subsequently, a minimum of two clones from each transfection were analysed for UV survival by colony-forming assays.

6.3 Results

In total, six deletion constructs were analysed to determine whether removal of parts of the gene would affect the ability of the *CSB* cDNA to correct the UV sensitive phenotype of UV61. Initially mixed populations of antibiotic resistant transfected cells were subjected to UV-selection as described in chapter 4. Of these six constructs, three failed to show any resistance above that of the untransfected cells, whereas the other three appeared to be more resistant.

6.3.1 C-terminal deletions

Two C-terminal deletions were created introducing stop codons at residues 1468 and 1383 that result in truncations of 26 and 111 amino acids respectively. Figure 6.1 shows the length of the deleted cDNAs with respect to full length and the helicase /ATPase domains. Two clones from each deletion were analysed for survival to determine the effects of the truncation and to ensure that there was no excessive clonal variation.

Figure 6.4 shows the survival curves generated for the clones carrying each of the deleted cDNAs. The survival curves in all cases are similar to that of the untransfected UV61 cells and considerably lower than the levels obtained after transfection with the wild-type. Therefore deletion of the C-terminus of the gene, even only 26 amino acids, destroys the DNA repair capacity of CSB.

6.3.2 N-terminal deletions

N-terminal deletions were created to produce cDNAs starting at amino acids 29 and 137, and the relative lengths of these cDNAs compared to the full length can be seen in figure 6.1. Two clones for each deletion were analysed for survival in a colony forming assay, the results of which are shown in figure 6.5. For comparison, the survivals for UV61, the parental AA8 and the wild-type corrected cells are also shown. Survival curves for each of the clones carrying N-terminally deleted cDNAs show intermediate levels of correction. The clones show levels of survival significantly higher than that of the mutant UV61, yet do not reach the levels of correction seen with the full length *CSB*. It is also interesting that the larger 137 amino acid deleted cDNA appears to restore more UV resistance to the cells than the 28 amino acid deletion.

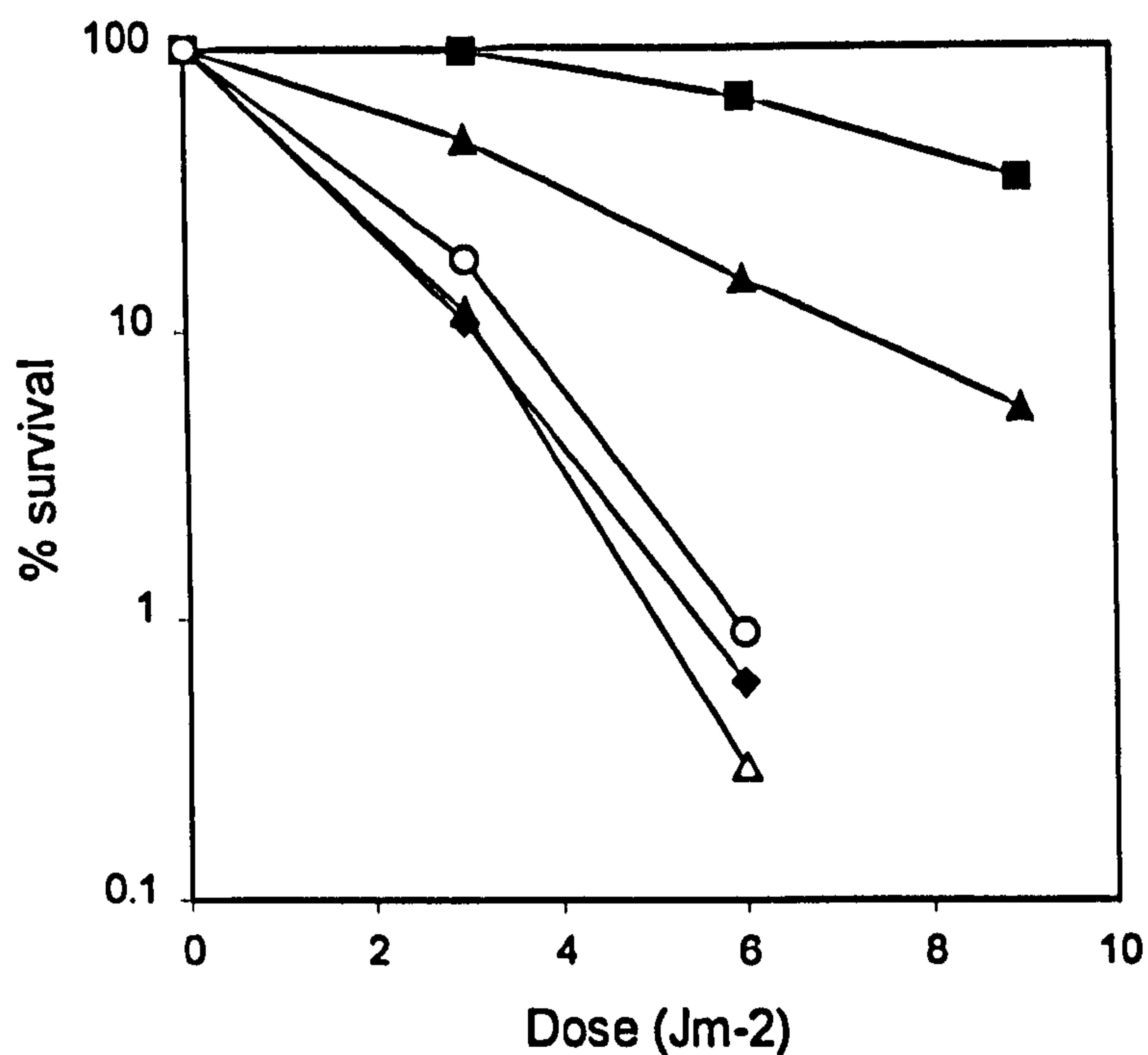


Figure 6.4 Survival curves for clones derived from transfection of UV sensitive UV61 cells with C-terminally truncated *CSB* cDNA. Cells transfected with cDNA missing either the C-terminal 25 (Δ) or 110 (\blacklozenge) amino acids fail to show any survival above that of untransfected UV61 cells (O). ■ = AA8, ▲ = WT corrected.

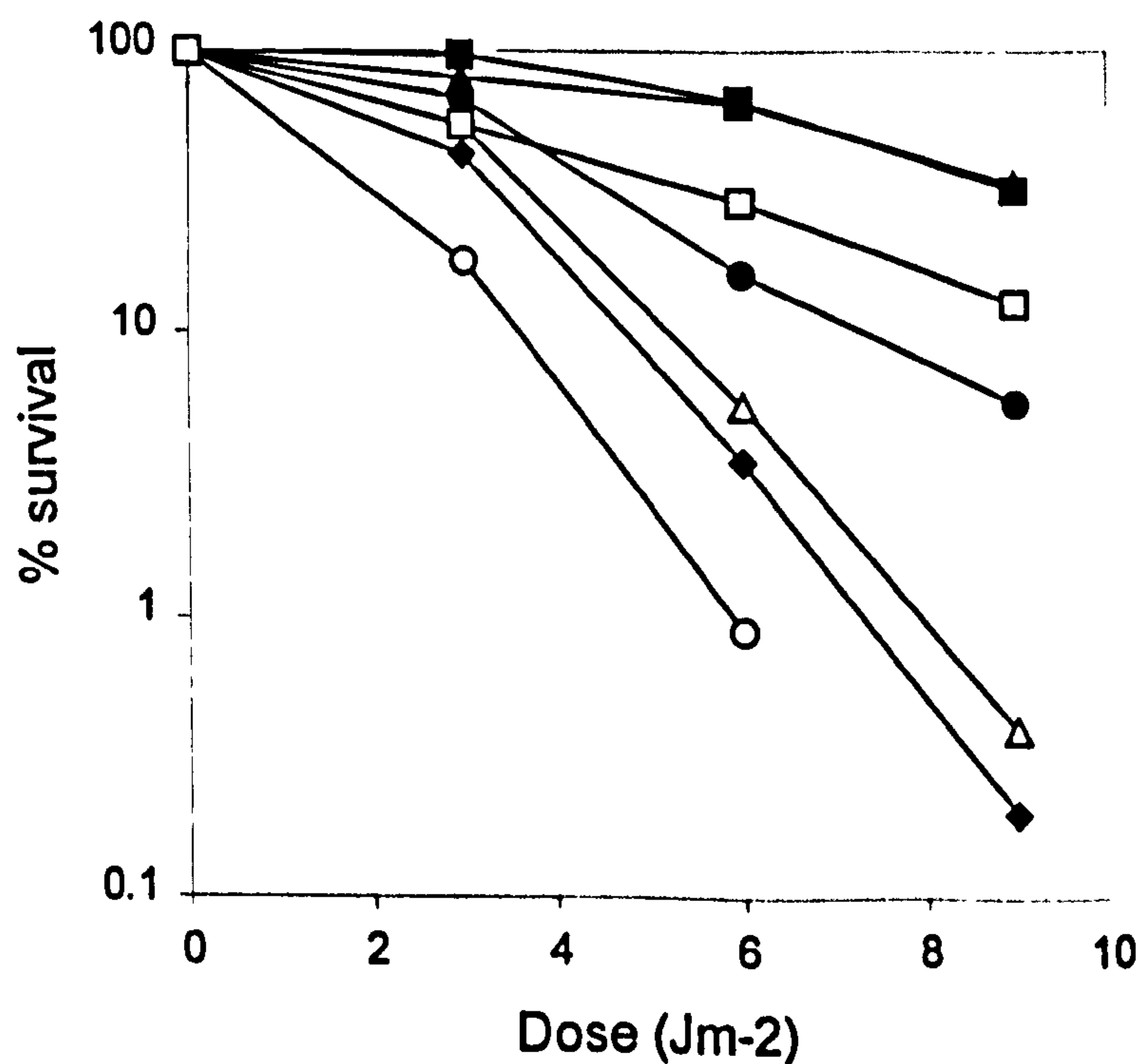


Figure 6.5 Survival curves for clones derived from transfection of UV sensitive UV61 cells with N-terminally truncated *CSB* cDNA. Cells transfected with cDNA missing either the N-terminal 25 or 137 amino acids show survival levels above that of untransfected UV61 cells (O), but do not restore the survival to that of WT corrected (▲) or parental AA8 cells (■). Δ = 25 (clone 1), \blacklozenge = 25 (clone 7), \bullet = 137 (clone 2), \square = 137 (clone 3).

6.3.3 Glycine and acidic domain deletions

Seven glycine residues and a glutamic acid residue positioned at amino acids 440 to 446, and seven glutamic acid residues at 378 to 384 were deleted from the cDNA. The position of these deletions relative to other domains within the gene can be seen in figure 6.1. The survival curves for the clones analysed from each transfection can be seen in Figures 6.6 and 6.7 along with the curves for AA8, UV61 and a wild-type transfectant. Removal of the glycine domain appears to destroy the DNA repair capacity of CSB. This could either be because of the functional significance of these residues, or just through changes caused by removal of a portion of the gene. In contrast, the deletion of seven glutamic acid residues from the acidic rich region seems to have a varying effect on the ability of the *CSB* cDNA to correct the UV sensitivity of UV61. A rapid survival screen of a mixed population of transfected cells suggested that the $\Delta 1211-1213$ cDNA was able to correct the UV sensitivity. However, survival curves on clones derived from the transfection showed an intermediate and variable sensitivity (Figure 6.7).

6.4 Discussion

Deletion analysis of CSB has revealed several interesting features of the protein. Removing even 26 amino acids from the C-terminus destroys all ability of the *CSB* cDNA to correct the UV sensitive phenotype of UV61.

The results are very different for the N-terminal end of the gene where substantial DNA repair activity is retained even when 137 amino acids are missing. This demonstrates that the N-terminal part of the protein is not essential for this activity of the CSB protein. Its necessity for other activities associated with CSB, transcription coupling or DNA dependent ATPase activity, is yet to be determined. Interestingly, the 137 amino acid deletion appears to correct better than the smaller 28 amino acid deletion. This may be due to the expression levels being higher in these clones. Alternatively, the structure may be

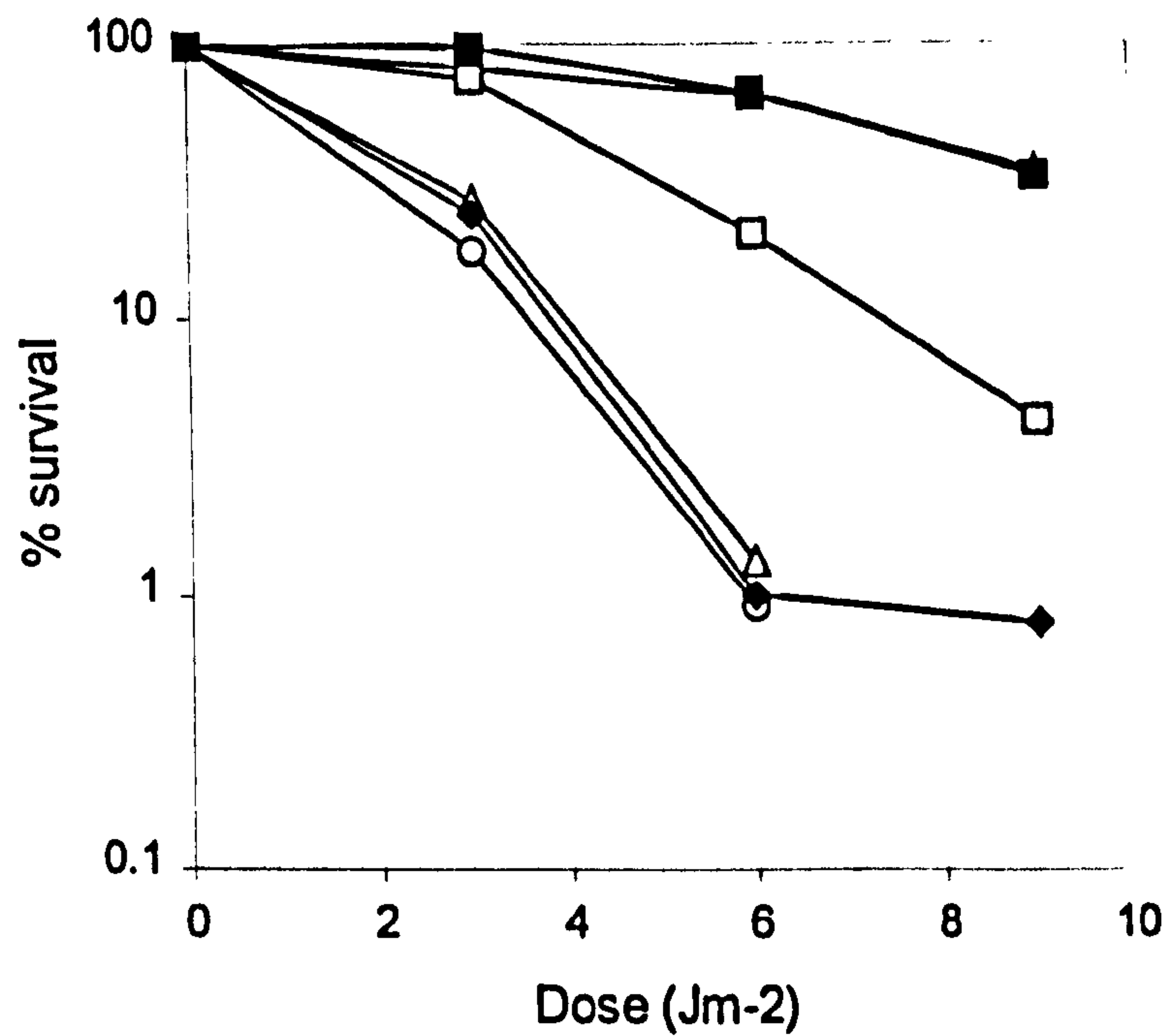


Figure 6.6 Survival curves for clones derived from transfection of UV sensitive UV61 cells with *CSB* cDNA deleted for the seven glutamic acid residues 378-384. Two clones (Δ and \diamond) fail to restore survival levels above that of untransfected UV61 cells (\circ). The third clone (\square) shows some increase in survival, but does not correct to the same extent as WT cDNA (\blacktriangle). ■ = AA8.

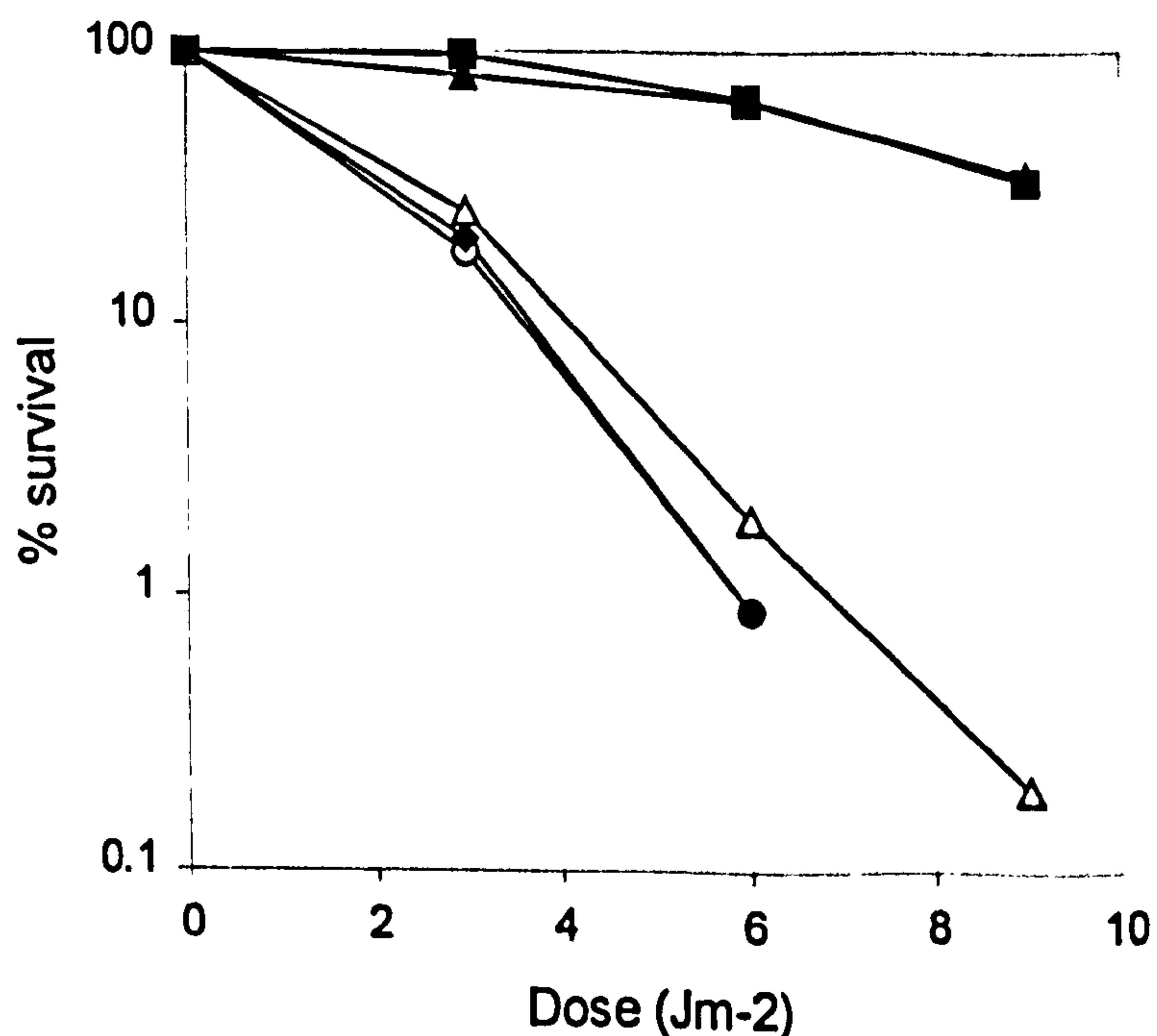


Figure 6.7 Survival curves for clones derived from transfection of UV sensitive UV61 cells with *CSB* cDNA deleted for the glycine domain at residues 440-446. The two clones shown (Δ and \diamond) fail to restore survival levels above that of untransfected UV61 cells (\circ). ■ = AA8, ▲ = WT corrected.

such that the smaller deletion leaves exposed or free tails that result in steric hindrance of protein folding or interactions.

The deletions of the glycine and acidic regions have given very surprising results. The removal of the glycine domain results in non-functional protein for the purposes of cell survival. Glycine is the smallest amino acid and has no side chains, often occurring in regions of flexibility. Therefore, removal of the seven consecutive residues could prevent the protein from adopting the correct conformation for activity or protein:protein interactions. Alternatively, it may be that removal of seven amino acids from the gene has in itself destroyed the protein structure irrespective of the residues deleted.

The removal of seven amino acids, in the case of glutamic acid residues from the acidic rich region, does not appear to have a particularly dramatic effect on the DNA repair function of CSB. The intermediate nature of the clones may be as a result of different expression levels. The variation between clones can be presumed to be a result of different protein levels, thus this may reflect an ability of the mutant protein to correct when over-expressed. Alternatively and less likely, this region may not be absolutely essential for the DNA repair activity of CSB and it is still able to carry out substantial repair. This is somewhat surprising as acidic regions are often involved in holding together protein structures, or are points of contact in protein:protein interactions. Furthermore removal of the seven residues will result in a drastic change in the overall charge of this part of the protein. Either the necessary interactions are being carried out by other means or domains, or sufficient acidic residues remain for the structure or complexes to hold together.

There is clearly a need for further investigation into functional domains and residues. The helicase/ATPase domains are good candidates for essential regions and mutational analysis of these motifs will be discussed in chapter 7.

CHAPTER 7 GENERAL DISCUSSION

Lowry et al., 1982 have classified CS into two groups. Type II is the more severe form of the disease with symptoms evident at birth and type I is milder with patients appearing normal at birth and symptoms becoming apparent during the first few years.

The aim of this project was to try and identify mutations in the CSB gene of patients with Cockayne syndrome. It was hoped that the mutations would correlate with the severity of the disease, with mutations resulting in severely truncated protein occurring in type II patients. Furthermore, a series of deletions were created to determine whether the N- or C-terminus of the protein was essential for function. Also, putative functional domains were deleted to see whether these regions were indeed functionally significant.

7.1 Mutational analysis

The full mutation spectrum identified in CS-B patients is shown in figure 7.1 (see also (Mallery et al., 1998)). Mutations resulting in severely truncated product and therefore probably no protein, have been identified in both alleles of patients indicating that the gene is not essential for viability. This is also the case in mice (van der Horst et al., 1997) and the yeast *Saccharomyces cerevisiae* (Verhage et al., 1996) where knocking out the respective CSB homologues does not affect viability. Interestingly, in both cases the phenotype is particularly mild when compared to that of CS-B patients. CS-B patients exhibit a wide variety of symptoms including UV sensitivity and neurological degeneration (see chapter 1 for more detail), and the severity of the disease is such that patients rarely survive beyond their teens. In contrast, CSB mice appear almost normal with typical CS symptoms occurring only very mildly.

There is no apparent correlation between the location or nature of the mutation and the clinical features, nor are there any particular mutational hotspots. Therefore the site and

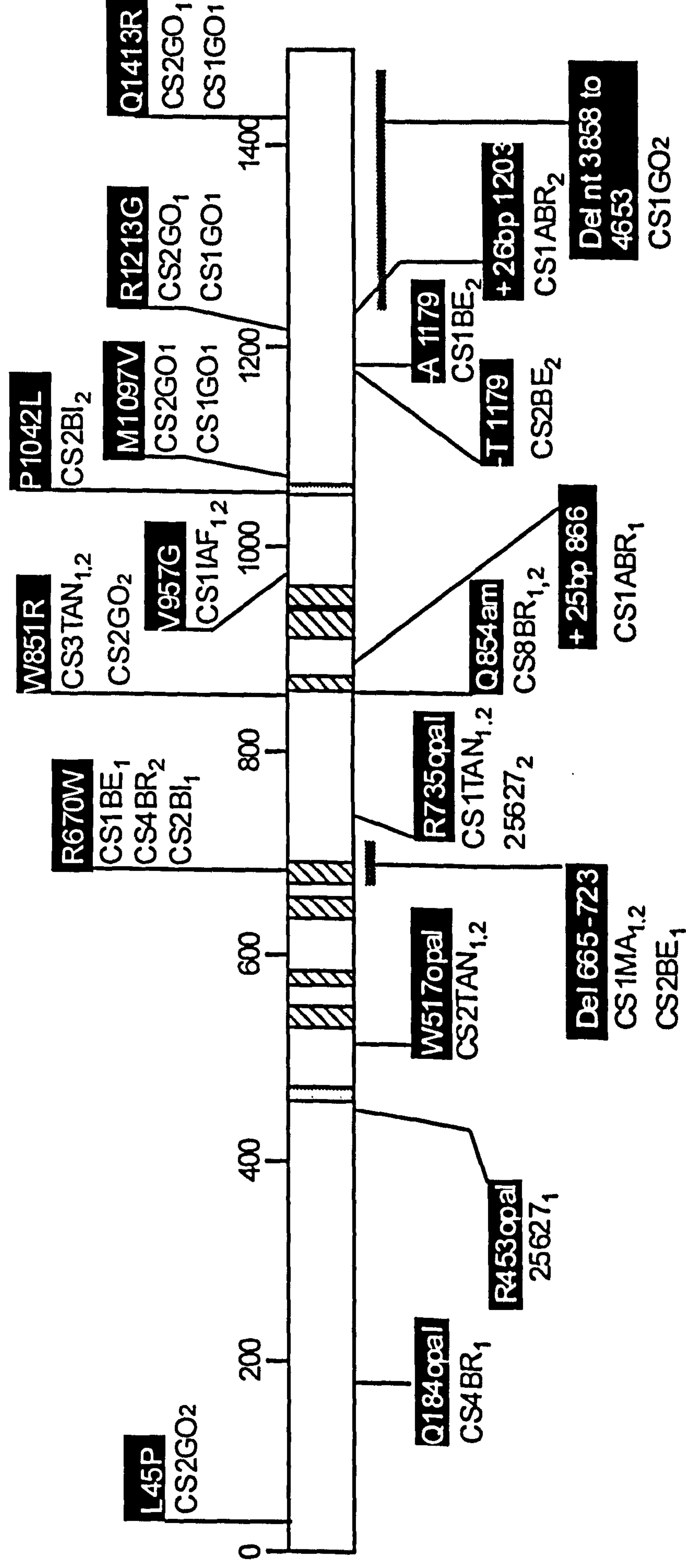


Figure 7.1 All mutations found in the CSB gene during this study of patients with Cockayne syndrome type B. Mutations shown below the gene are inactivating, and above amino acid substitutions. The black boxes show the mutation with the patient designation below and alleles are indicated in subscript. \square = helicase domains, \square = nuclear localisation signal.

type of mutation is not sufficient to explain the variation in clinical features. Cockayne syndrome is a DNA repair disorder and as such any residual non-repaired damage might have an influence on the health of the cells of CS patients. Cells from patients with CS are deficient in the repair of actively transcribed DNA, and it is likely that damage will persist within the life-cycle of these cells leading to mutations in other genes. CS cells are also known to have an enhanced mutation frequency in response to UV (Arlett and Harcourt, 1983). Levels of repair are particularly significant when considering cell types such as neurones that are not generally renewed and have a long life span. Cell death as a result of non-repaired damage in these cells will have significant phenotypic consequences, although the neurological defect in CS patients is dismyelination of neurones rather than neurone death as is found in complementation groups A and D of XP. Consequently, greater exposure to DNA damaging agents may result in a more severe phenotype.

A major difference between clinical features in CS patients is the time of onset of the disease and the speed of deterioration and there is also great variation in the types of symptoms observed. Under these circumstances it is reasonable to predict that the environment might influence the stage in development at which the disease becomes apparent. For example, the levels of nutrients reaching a foetus may have an effect, since a healthy foetus may be better able to tolerate a lack of or mutant CSB. Similarly, if there is substantial exposure to DNA damaging agents at an early stage of development, it might be expected that the symptoms would be more severe and evident at birth. Under both these circumstances, the health of the mother, either through the environment, nutrition or through smoking might affect the general health of the CS child and possibly influence the severity of the resulting disease.

In addition to the environmental effects described above, the genetic background of the patient may affect the disease severity and symptoms. For example, the efficiency of global genome repair or alternative repair pathways may vary between individuals.

Therefore the levels of damage persisting is also likely to vary, thereby influencing the resulting phenotype. Other genetic factors may also enable or prevent the CS patients from tolerating or preventing damage. There may be genes not related to the repair pathways that will allow or prevent transcription and replication to bypass DNA lesions under certain conditions. Therefore, mutation of these genes might upset the balance of the cell and create a situation whereby there is too much or little damage remaining in the genome.

Interestingly there are two siblings, GM10903 and GM10905 who present a classical XP phenotype, but have been shown to belong to the CS-B complementation group (Itoh et al., 1996). The mutation present in these patients is the same as those found in the patients CS1TAN and in one allele of 25627 from this study (Figure 7.1) who have a classic CS type I phenotype (S. Colella, M. Stefanini, personal communication). An identical mutation can therefore result in a wide variety of clinical features, with effects extending beyond the severity of the CS symptoms. In this case it is very difficult to predict what has caused this improbable phenotype, since insufficient information is available about the exact function of CSB and the XP gene products during development.

XP-G patients with XP alone and XP combined with CS have been studied with reference to their involvement in the transcription-coupled repair of oxidative damage (Cooper et al., 1997). Only cells from patients exhibiting CS symptoms are defective in the TCR of thymine glycols. Therefore, a defect in preferential repair of oxidative damage correlates with the presence of the CS phenotype in XP-G patients. It is possible that persistence of un-repaired oxidative damage may result in some of the symptoms not related to a NER defect. Further investigation into the cause of XP with CS rather than XP alone in patients, when the mutations occur in the same gene may also shed light on the genotype phenotype relationship of CS-B.

Although the mutation spectrum did not show any hotspots in the gene or any correlation with the clinical features, it did suggest a number of possibilities as to important domains, and regions that may not be absolutely necessary for function.

The leucine to proline mutation at amino acid 45 does not affect the ability of the protein to correct the UV sensitive phenotype of UV61 CHO cells. A leucine to proline change is likely to result in conformational changes due to the unique nature of the amino acid. For example if this region of the protein is α -helical in structure the disruption is likely to be greater since proline is the only amino acid whose side chain does not project out from the helix. Therefore, the ring structure of the proline side chain will cause steric hindrance to the α -helical structure and will usually result in a significant bend in the helix. It can be predicted that the N-terminus of the protein can be disrupted at least to some extent without loss of function. In contrast, the 3' end of the gene appears to be essential, since the deletion of the C-terminal 233 amino acids in patient CS3GO results in a complete loss of function, as measured by its ability to complement the UV sensitivity of UV61. The majority of the mutations and more importantly the missense mutations are located in the C-terminal two-thirds of the gene, suggesting that this region and particularly the central region may be functionally the most important. This is also the region in which the putative functional ATPase, glycine, and acidic rich domains are located.

Such observations led to the work on deletion analysis of the gene to study the effects of deleting both the C- and N-termini of the protein, and two putative functional domains, an acidic rich region and a stretch of glycine residues.

In agreement with the prediction from the C-terminal truncation in the patient CS3GO, deletion of the C-terminal 26 amino acids of CSB totally destroys the ability of the cDNA to complement UV61. Also, the data obtained from the deletion of the N-terminal end of CSB cDNA supports the idea that deleted cDNA are able to at least, partially complement the UV sensitivity of UV61 cells. Clones carrying an intact cDNA as judged by PCR

correct to some extent, but there is considerable variation between clones, with deletion of the 488 N-terminal residues resulting in greater levels of survival than clones missing the N-terminal 154 residues. There are a number of possible explanations both for the larger deletion correcting to a greater extent than the smaller, and for the clonal variation.

Firstly, the structure of the CSB protein may be such that deletion of a smaller part of the N-terminus disrupts the protein structure. This disruption may result in a free tail that could cause steric hindrance to protein:protein or protein:DNA interactions. Since the ATPase activity of CSB is dependent on DNA (Citterio et al., 1998; Selby and Sancar, 1997) it is likely that a good interaction is necessary for enzymatic activity. Similarly, CSB resides within a large molecular weight complex (van Gool et al., 1997) and any disruption of interactions between proteins of the complex is likely to at least, partially, disrupt the activity of the complex in some or all of its functions. Therefore, if the deletion affects such interactions, this could explain why the larger deletion might correct more effectively, since the free tail would be missing or shorter.

An alternative reason for the differences in levels of survival in the clones bearing N-terminal deletions may be differences in levels of protein expression. If the levels of CSB are regulated in some way, it is possible that this regulation might be facilitated by the N-terminus of the protein. Alternatively the expression levels obtained from this vector may be variable and the levels may be such that interactions or activities that occur may be dependent on the amount of protein present.

With either of the above, it is possible that higher levels of expression might be able to counteract the effects of the deletions. The N-terminal deletions may be sub-optimal in their ability to form interactions, and therefore greater protein levels may increase the chances of such interactions occurring.

Further mutational analysis has investigated the functional significance of the glycine and acid rich regions within the gene. The region of seven consecutive glycine residues

appears necessary for function, since deletion of these residues results in a cDNA that is unable to complement the UV sensitivity of UV61 (Figure 6.6). However, several interpretations of these data are possible. For example, deletion of seven amino acids from the protein might result in the three dimensional structure of the protein being disrupted. Furthermore, glycine residues are the most flexible amino acids and allow unusual conformations through the lack of a side chain. It may therefore be possible that a change in these residues from the original glycine might alter the protein structure considerably. Hence, the deletion may well result in the disruption of a functional protein structure rather than of an activity, either through the loss of the glycine residues or seven amino acids. The same may be true for the deletion of seven glutamic acid residues from the acidic rich region, although the inactivating nature of this deletion is less clear (Figure 6.7). There appears to be some level of correction from the truncated protein, and this suggests that the secondary structure of the protein is not completely disrupted. The acidic domain stretches over thirty-nine amino acids, of which twenty-two are either glutamic acid or aspartic acid residues (Figure 6.3). The functional activity partially remains with respect to correction of the UV sensitivity of UV61, suggesting that the remaining acidic residues are at least to some extent able to carry out the necessary role required for NER. Since acidic residues are commonly found in regions of interaction between proteins, it is feasible that a smaller acidic domain when over-expressed is capable of forming such interactions, albeit less efficiently.

7.2 Western blots

To determine the inactivating nature of base substitution mutations found in CS-B patients, the changes were created in an expression vector and transfected into UV61 hamster cells as described in chapter 4. The presence of the transfected cDNA in the genome of the recipient cell line was determined by PCR from genomic DNA. However, presence of the

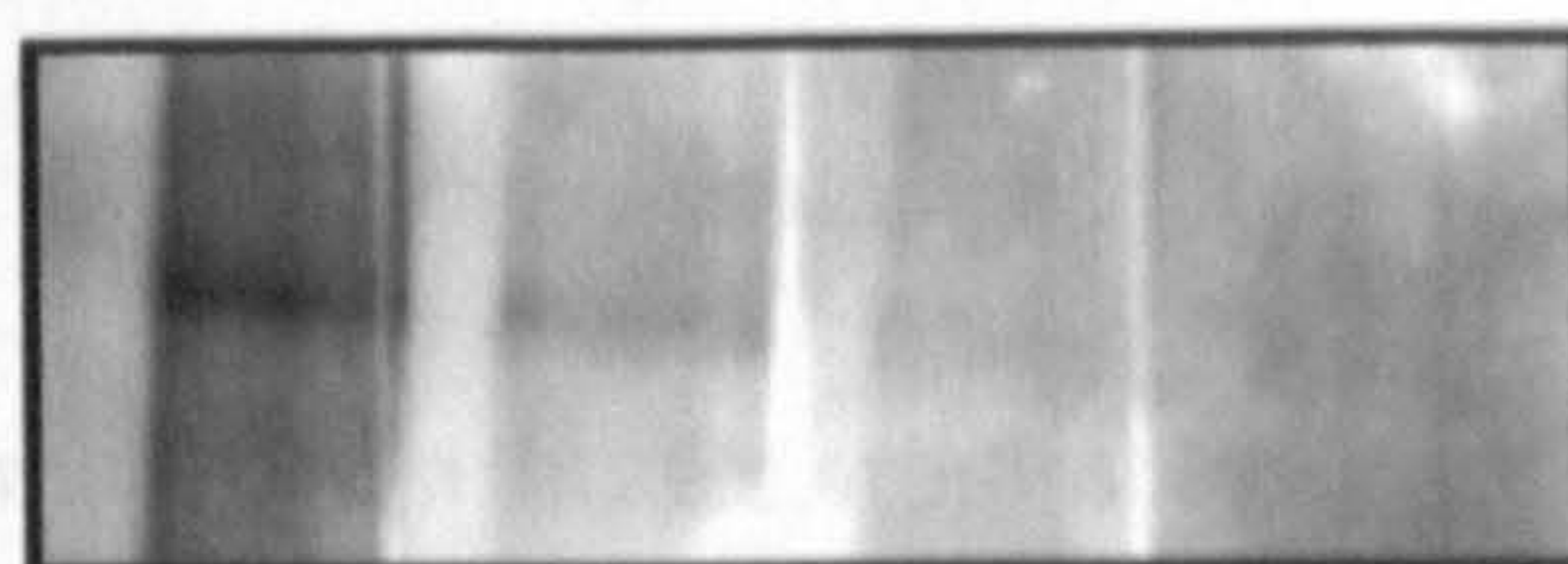
DNA does not necessarily result in expression of the protein either because of rearrangement in the transfected vector or rapid degradation of protein produced, particularly if it is mutant. The results of several transfections with mutant cDNAs have shown considerable clonal variation in the ability to rescue the UV sensitivity of UV61. As discussed earlier there are several theories that may explain such variation, including differences in the levels of human CSB protein expressed. To determine whether this was the case and to confirm that transfected cDNAs were being expressed in the hamster cells, Western blots were carried out.

Early attempts were carried out using a CSB antibody raised against a C-terminal peptide (van Gool et al., 1997) and detected using the ECL system. Later, an alkaline phosphatase conjugated secondary antibody was used for detection of CSB antibody bound to the filter. Efforts were also made to detect protein in cells transfected with HA tagged cDNA constructs using a monoclonal HA antibody. Several different methods have also been employed for preparation of extracts including cell fractionation, freeze-thawing, detergent lysis, sonication in PBS and boiling in SDS sample buffer.

Early experiments failed to show any significant bands of around the correct size, and in the absence of a positive control, it was difficult to determine whether any bands seen were in fact CSB. A positive control was obtained through expression experiments in the baculovirus system, with a partially purified tagged recombinant protein giving a very strong signal on Western blots (see section 7.3.2). Decreasing amounts of protein could then be used to determine the sensitivity of the different antibodies and detection systems. With ECL detection, monoclonal HA antibodies (Figure 7.2a), either commercially obtained or purified from 12CA5 hybridomas (Figure 7.2b) were more sensitive than the CSB specific antibody. However, when using alkaline phosphatase detection, the sensitivity of CSB was equal to that of the HA (Figure 7.2c). Since many of the earlier transfections were performed using non-tagged constructs, it was decided to proceed with

Volume
protein(μ l)

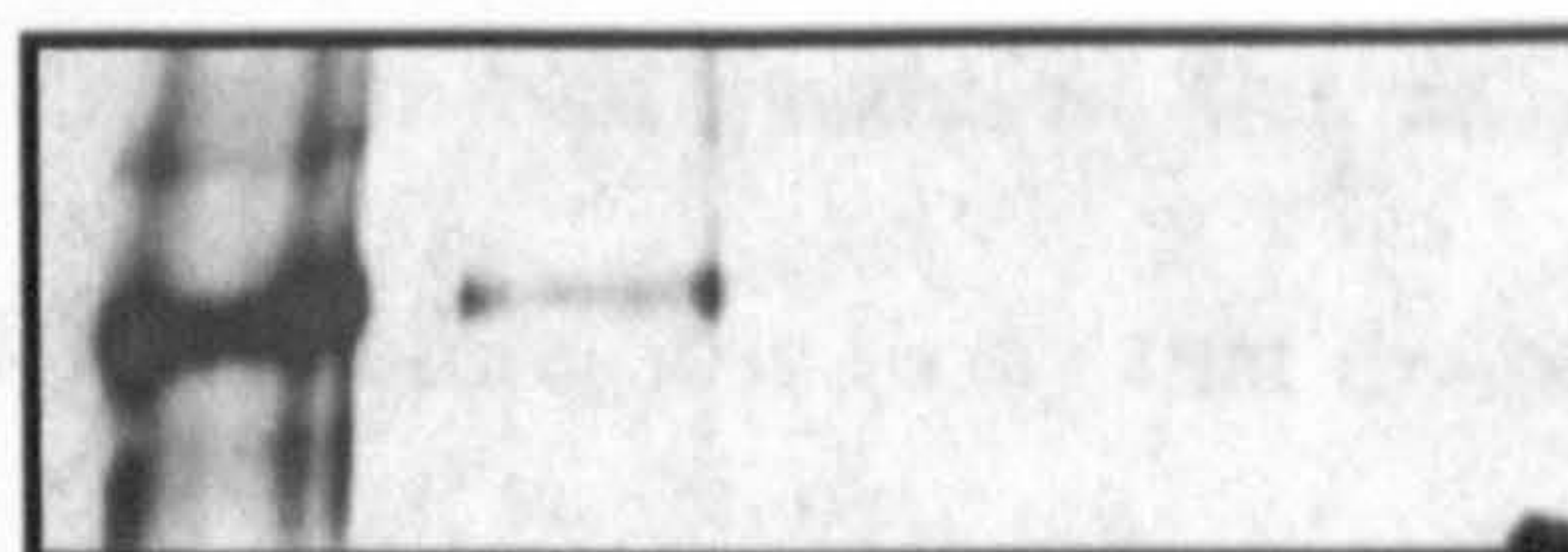
20 5 2 1



a

Volume
protein(μ l)

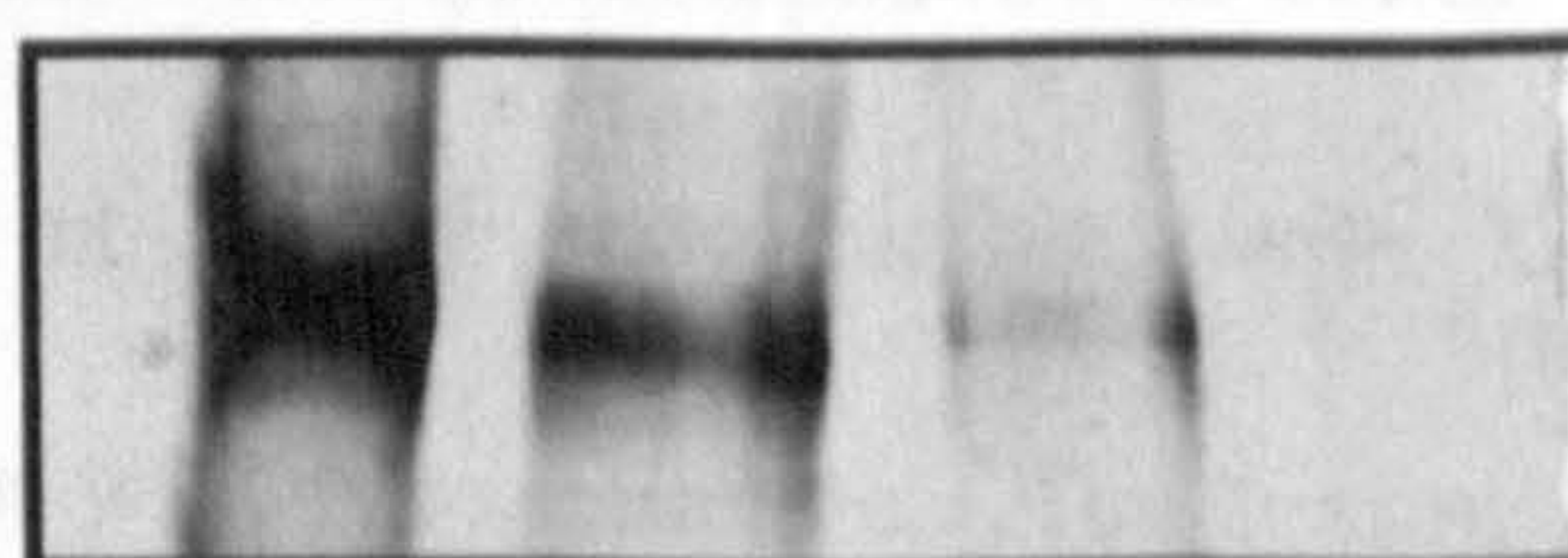
10 5 2 1



b

Volume
protein(μ l)

20 10 5 2



c

Figure 7.2 Western blots of CSB positive control (obtained from expression of recombinant protein through baculovirus) to test the sensitivity of a commercially bought monoclonal HA antibody (panel a), monoclonal HA purified from 12CA5 hybridomas (panel b) and polyclonal CSB antibody (panel c). All three antibodies produce a good signal with 5 microlitres of protein.

the CSB antibody. Whole cell extracts of corrected hamster cell lines failed to produce a definite CSB band, with the extracts made either by sonication in PBS or freeze-thawing. In order to concentrate the CSB protein in proportion to other proteins in the cell, nuclear preparations and extracts were made. Again, these extracts failed to show CSB specific bands.

Whole cell extracts were made in a buffer, routinely used as a first step in immunoprecipitations, containing a cocktail of protease inhibitors and detergent to lyse the cells and nuclei. This time there were bands, which were slightly smaller than the positive control in lanes for extracts from cells transfected with both wild-type and mutant cDNAs (Figure 7.3). However, when extracts were made by the same method for non-transfected UV61 cells, the resulting Western blot produced the same band as present in the transfected cells (Figure 7.4). Therefore, it is likely that the band seen in these extracts is either hamster CSB or a non-specific protein that is recognised by the antibody.

To check whether the lack of detection of human CSB in the transfected UV61 cells was a detection or blotting problem, a blot containing human nuclear extracts was obtained to try and detect the human protein in human cells. Native levels of human CSB were easily visible with CSB antibody and alkaline phosphatase detection (Figure 7.5). This blotting and extraction procedure is highly reproducible, with CSB detectable in several independent extracts from a wild-type cell line. The reason for the lack of detection of human CSB over-expressed in hamster cells is not clear, but it severely hindered interpretations of some of the transfection experiments.

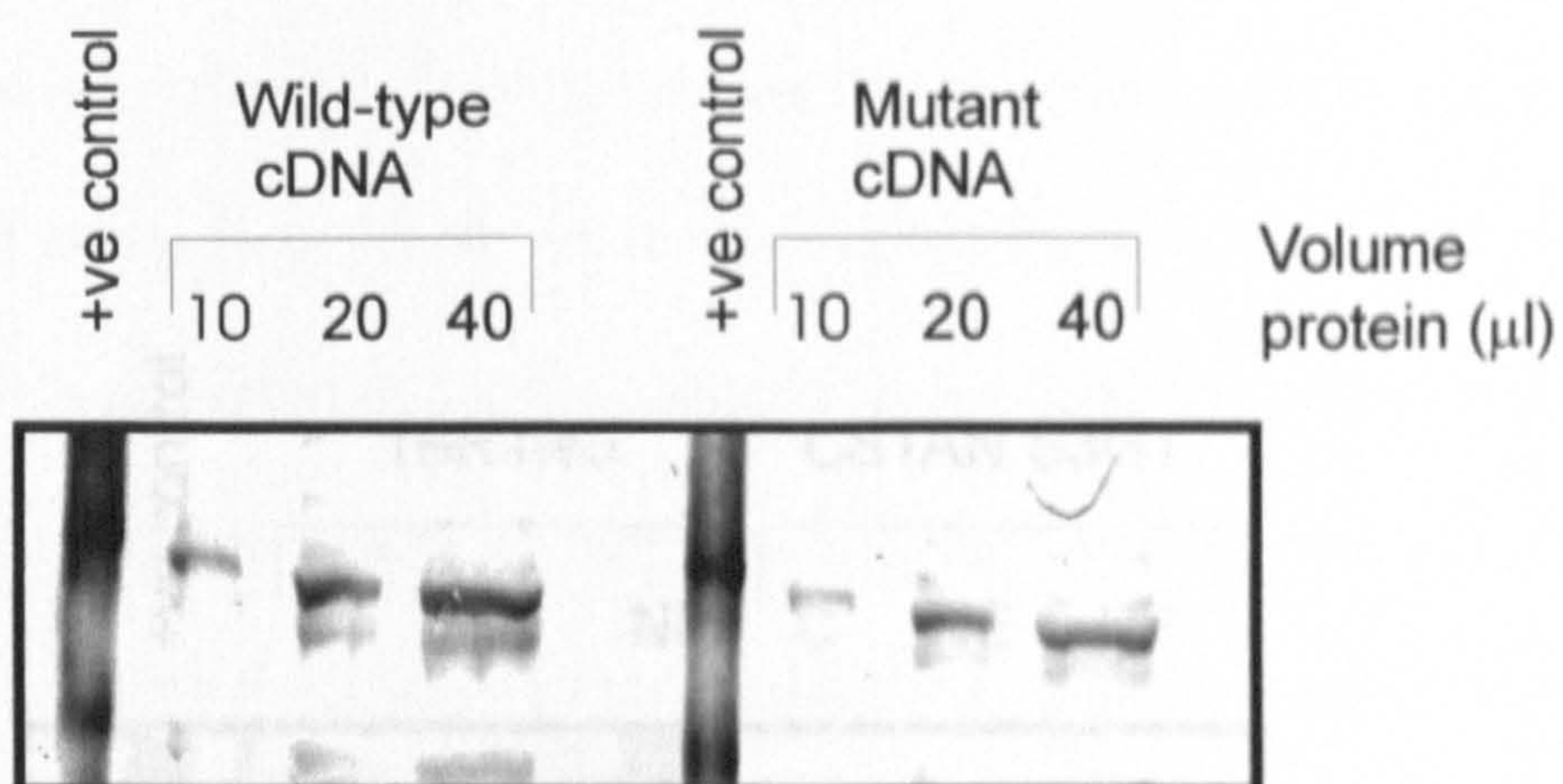


Figure 7.3 Western blot with varying amounts of cell extracts from cells transfected with a wild-type or mutant *CSB* cDNA. Detection was with polyclonal *CSB* antibody and alkaline phosphatase.

Figure 7.3. Western blot of cell fractionation of cells wild-type and transfected with mutant *CSB* cDNA, demonstrating that *CSB* is detectable at active levels in transfected cells. The lack of protein in the cells null for *CSB* indicates that the gene product for the null cells is *CSB*.

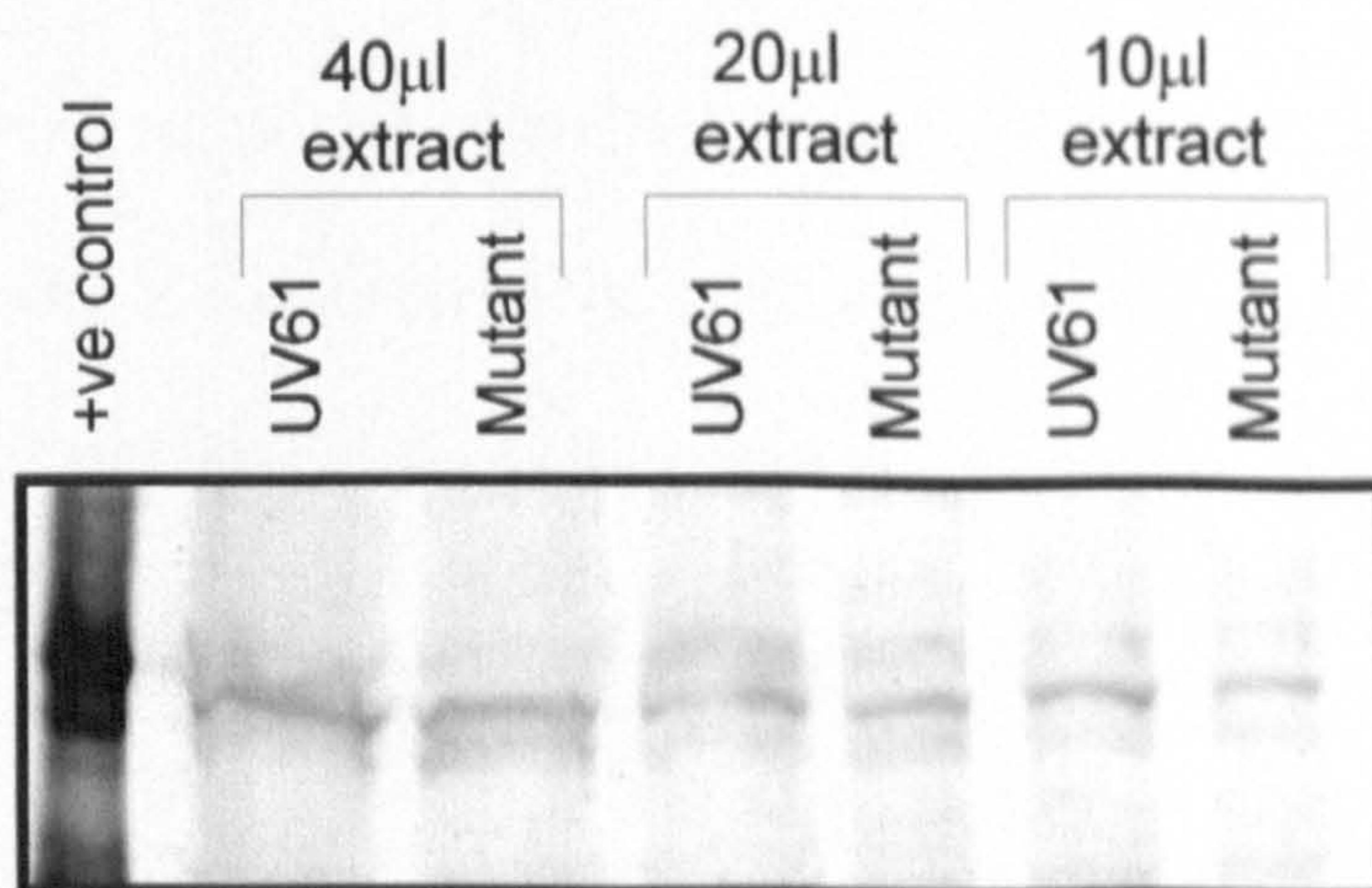


Figure 7.4 Western blot with varying amounts of cell extract from untransfected UV61 cells and cells transfected with a mutant *CSB* cDNA.

7.1 Future Work

7.3.1 Functional analysis of the wild-type glycosylase

As discussed in section 1.1, deletion of a number of coding exons in *CSB* results in

loss of function, irrespective of the functional significance of the 3' region.

Therefore, it is important to determine whether the information gained from the

analysis of the wild-type glycosylase is sufficient to explain the phenotype of the

mutant. This is proposed to be done by analysing the expression of the wild-type

protein in the cells null for *CSB* and comparing the results with those of the

mutant. This will be done by analysing the expression of the wild-type protein

in the cells null for *CSB* and comparing the results with those of the mutant.

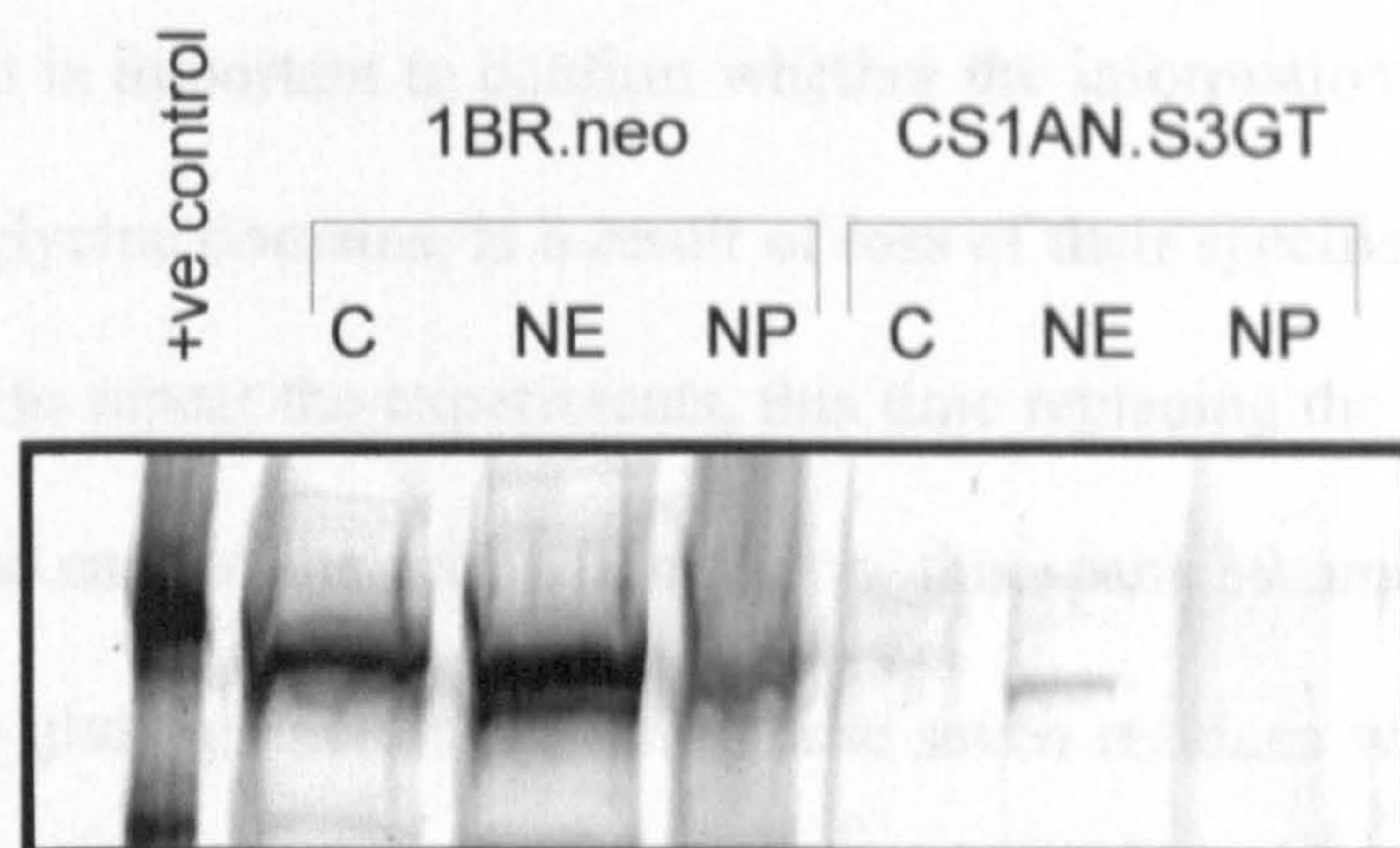


Figure 7.5 Western blot of cell fractionation of cells wild-type and mutant for *CSB*, demonstrating that *CSB* is detectable at native levels in transformed human fibroblasts. The lack of protein in the cells null for *CSB* indicates that the band present for the wild-type cells is *CSB*.

with glucose should only change the ratio of the two bands.

The glycosylase domain is a stretch of seven consecutive glycosylase domains that are

region of flexibility within the protein structure. The first is located in the

region of flexibility within the protein structure. The first is located in the

region of flexibility within the protein structure. The first is located in the

region of flexibility within the protein structure. The first is located in the

region of flexibility within the protein structure.

In both cases, the control will be used to determine the level of the

glycosylase domain in the CSB protein. Furthermore, the results will be

used to determine the level of the glycosylase domain in the CSB protein.

These and other results will be used to determine the level of the

glycosylase domain in the CSB protein.

7.3 Future Work

7.3.1 Functional analysis of the acidic and glycine domains

As discussed in section 7.1, deletion of a number of amino acids from a protein may result in loss of function, irrespective of the functional significance of the deleted fragment. Therefore, it is important to confirm whether the information gained from deletion of the acidic and glycine domains, is a result of loss of their specific function. For this reason it is proposed to repeat the experiments, this time replacing the residues rather than deleting them. In the case of the acidic rich region, this spans 39 amino acids and contains seven consecutive glutamic acid residues. These seven residues will be replaced by glutamine residues. This should not disrupt the overall protein structure of CSB through steric hindrance, since the structure of glutamine is virtually identical to that of glutamic acid, with the acidic group being replaced by its amide. Therefore, replacement of glutamic acid with glutamine should only change the acidic nature of the seven residues.

The glycine domain is a stretch of seven consecutive glycine residues that may give a region of flexibility within the protein structure. The idea of future experiments is to replace alternate residues with serines. Serine is also a small amino acid and like glycine is un-charged, and as such is unlikely to cause any steric hindrance or disruption of structure. This will determine whether the presence of the glycine residues in this region is functionally significant.

In both cases, the mutant cDNAs will be tested for functional complementation of the UV sensitive phenotype of the CHO mutant UV61. Furthermore, recombinant protein derived from the deletion constructs and the corresponding mutant constructs will be tested for ATPase and DNA binding activity through expression and purification from the baculovirus expression system (see below).

7.3.2 Expression and purification of recombinant proteins

The baculovirus expression system allows the production of high levels of recombinant protein. The main advantages of this system over expression in *E.coli* or yeast is that mammalian proteins are more likely to be soluble and the insect cells in which the proteins are expressed are able to carry out a similar level of post-translational modification to that seen in mammalian cells.

For the purposes of the CSB protein, the BAC-TO-BAC baculovirus expression system from GibcoBRL is being used with infection of *Spodoptera frugiperda* (Sf9) cells under serum-free conditions in suspension culture. The CSB cDNA is expressed in the pFastBac vector and possesses two fusion tags, a His tag at the 3' end and HA at the 5' end. The protein will be purified in three stages as determined by van Gool et. al., 1997. The first step of purification takes advantage of the DNA binding capacity of CSB by running cell extract through a heparin column to eliminate proteins that do not bind DNA from the sample. The fractions eluted from the heparin column are then purified on a nickel column that binds the His tag. The partially purified CSB is finally bound to monoclonal HA antibody and run down a protein G column. The protein G resin is able to bind immunoglobulins and will therefore bind the HA antibody now attached to CSB. Purified CSB can be eluted from the column with an HA peptide. Purified CSB will be analysed for DNA-dependent ATPase activity and DNA binding activity, in an attempt to identify regions of the protein required for these functions.

7.3.3 SNF2-like mutations

As described in chapter 1, CSB is a member of the SWI/SNF family of DNA-dependent ATPases, and as such possesses the seven characteristic domains. Peterson et. al. created a series of mutations around the ATPase domains of *S.cerevisiae* SNF2, some resulting in loss of activity and others dominant negative activity (for more detail see Richmond and

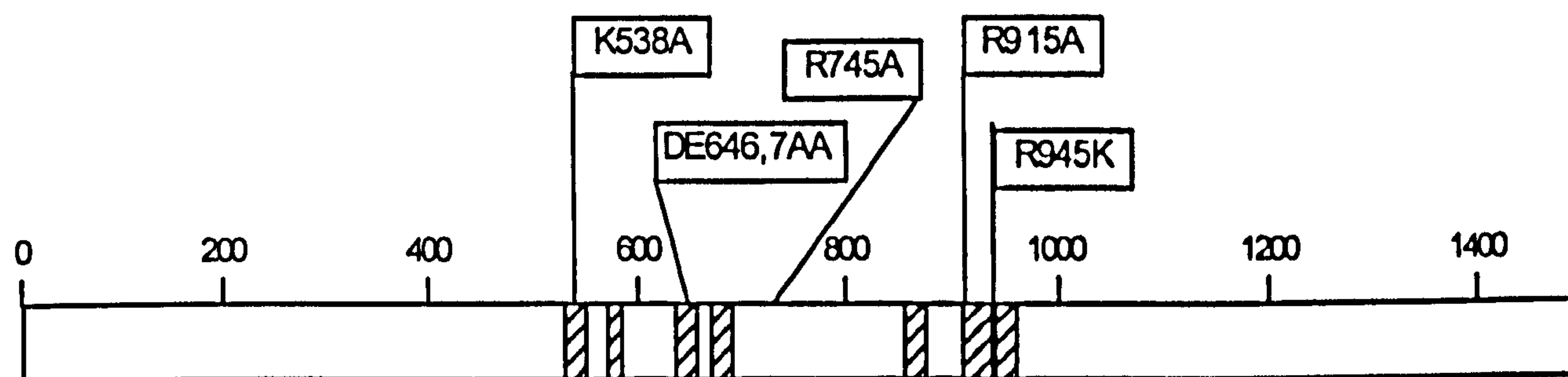


Figure 7.6 Diagrammatic representation of CSB showing the position and nature of the SNF2-like mutations created for functional complementation studies and protein assays.
 ▨ = helicase domains

Mutation in CSB	Mutation in SNF2	Phenotype in SNF2
K 538 A	K 798 A	+ Dominant negative
DE 646,7 AA	DE 894,5 AA	No activity
R 745 A	R 994 A	++ Dominant negative
R 915 A	R 1164 A	Low activity
R 945 K	R 1196 K	++ Dominant negative

Table 7.1 The nature of the mutations created in CSB and the analagous mutations in SNF2 from which they were derived. Also shown is the phenotype observed in the SNF2 mutants in yeast.

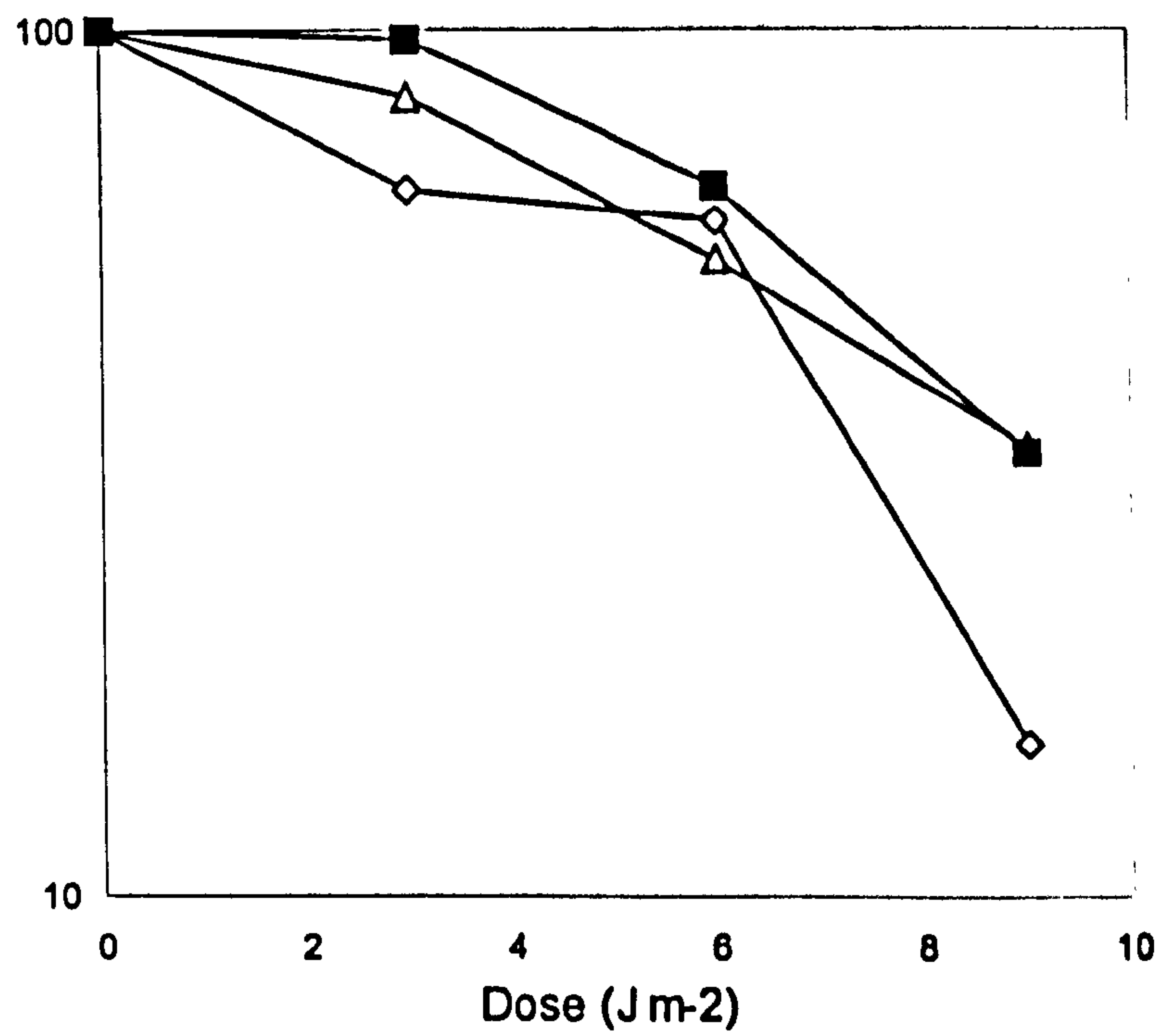


Figure 7.7 Survival curves showing the effect transfection of CSB mutant cDNAs has on the repair ability of wild-type AA8 cells. The cDNA deleted for seven glutamic acid residues has no effect (Δ) on AA8 survival, whereas the K538A mutation results in survival levels (\diamond) just below those of AA8 (\blacksquare).

Peterson, 1996). Analogous mutations have been created in the *CSB* cDNA, the mutations and the respective *SNF2* mutations are shown in Figure 7.6 and Table 7.1. The mutant cDNAs created will initially be transfected into UV61 and AA8 cells to determine their ability to correct or induce UV sensitivity in these cell lines.

The mutations K538A and R745A have been transfected into AA8 to determine potential dominant negative activity. Preliminary data indicates that K538A results in some dominant negative activity with the UV survival showing levels consistently below that of untransfected AA8 cells (figure 7.7). In contrast the second mutant transfected, R745A has no effect on the ability of AA8 cells to withstand UV irradiation. A similar response for both mutations when in *SNF2* is seen in *S.cerevisiae* (Richmond and Peterson, 1996).

Dominant negative activity is commonly seen in proteins that function as part of a complex, where the mutant protein competes with the wild-type to form the complex and any that possess the mutant will be non-functional. The dominant negative activity could also result from the complex or CSB alone blocking the access of the repair enzymes to a lesion, since it has been predicted that the ATPase activity may help to move a stalled RNA Polymerase from the lesion site (see below).

To further characterise the nature of the mutant protein and determine how the dominant negative effect might come about, the proteins will be purified and tested for ATPase and DNA binding activity.

7.4 Models for the Role of CSB

There are a number of different models proposed for the function of CSB in transcription-coupled repair. As discussed in chapter 1, human cells lacking functional CSB are unable to repair lesions in the transcribed strand of active genes. Therefore the protein is clearly required for functional TCR and as such is likely to have some involvement in transcription itself.

CSB has been linked with the movement or removal of a blocked RNA polymerase molecule and the subsequent recruitment of the repair proteins (Hanawalt, 1994). On encountering a lesion, RNAPII stalls (Hanawalt, 1994) and therefore may be required to retract or dissociate to allow repair enzymes access to the lesion (Donahue et al., 1994; Selby and Sancar, 1990). Transcription pause sites occur frequently and in *E.coli*, these result in backtracking of the polymerase molecule. This is accompanied by a shortening of the transcript, followed by resumption of RNA synthesis, in an attempt to bypass the pause site (Kassavetis and Geiduschek, 1993). It is possible that backtracking occurs in an attempt to bypass the stalling, with dissociation occurring only when this fails. Therefore it has been suggested that CSB could be involved in the backtracking or dissociation of RNAPII (Hanawalt, 1994; Troelstra et al., 1992; van Gool et al., 1997). However, it has been shown *in vitro* that neither recombinant CSB alone, nor with CSA, is able to prevent a stalled RNAPII molecule from resuming transcription, hence leading to the conclusion that these proteins cannot dissociate stalled RNAPII (Selby and Sancar, 1997). This may not disprove a lack of involvement of CSB in this process, since dissociation may be an indirect effect of the protein requiring other factors not present in this assay. Also, this experiment does not test for the ability of the CS proteins to induce backtracking of the polymerase, which may act as a signal for dissociation or ubiquitination (see below).

The model initially accepted by most groups as to the role of CSB was proposed by comparisons with the transcription-coupled repair pathway and MFD in *E.coli*. The *E.coli* MFD protein is also known as the transcription-repair coupling factor, and as such facilitates the preferential repair of the transcribed strand of active genes. TRCF acts by removing a RNAP molecule stalled at a lesion and attracting the NER proteins to the damage site (see chapter 1 for more detail). Whether CSB alone or along with CSA acts as a TRCF in mammalian cells is the subject of debate, although little data exists. Work with the *S.cerevisiae* CSB homologue Rad26 has suggested a function more like an uncoupling

factor, by allowing the basal transcription factor TFIID to return to transcription mode after repair (van Gool et al., 1994). This has also been suggested for human CSB (van Oosterwijk et al., 1996). CS cells are hypersensitive to the DNA damaging agent NA-AAF and are unable to recover RNA synthesis after exposure to this agent. The lesions caused by NA-AAF are repaired as efficiently in CS cells as in wild-type cells, and without strand bias. Therefore the recovery of RNA synthesis and repair of these lesions are clearly not linked, suggesting that TFIID is trapped in repair mode and that the CS gene products are required for converting it back to transcription mode (van Oosterwijk et al., 1996).

CSB is a member of the SWI/SNF superfamily of ATPases and helicases, many of which are involved in chromatin remodelling (see chapter 1). Like many of the other family members CSB resides in a high molecular weight complex, in this case of over 700kDa (van Gool et al., 1997) and possesses a DNA-dependent ATPase activity that is preferentially stimulated by dsDNA, either naked or nucleosomal (Citterio et al., 1998). It has been suggested that members of the SWI/SNF family are able to move along DNA, consequently destabilising bound proteins using energy derived from ATP hydrolysis (Pazin and Kadonaga, 1997). On the basis of such an activity, it was suggested that CSB may facilitate resumption of transcription after repair or pausing, through dissociation or destabilisation of associations between NER factors or nucleosomes and DNA (Citterio et al., 1998). This theory is not readily compatible with the observation that CS cells appear to possess a looser chromatin structure, in which chromatin-associated proteins are not as tightly bound as in normal cells. The experiments that instigated this theory measured the percentage of proteins released from the DNA after treatment with detergent and the resulting rates of transcription of the extracts. (Balajee et al., 1997). There was however no analysis of the additional proteins released in CS cells compared to normal cells, hence it is not clear whether these proteins are chromatin or transcription related. Also, the end-point of these experiments measured transcription rate and as described below there are a number

of potential problems associated with such a measurement in CS cells. Therefore, it is not certain that the predicted "loose" chromatin structure is significant to the phenotype of CS cells.

Furthermore, CSB has been shown to interact *in vitro* with RNAPII (Tantin et al., 1997; van Gool et al., 1997) and the ATPase activity associated with CSB is important for this interaction (Tantin et al., 1997). Selby and Sancar have reported that CSB enhances RNAPII elongation in a mechanism similar to the elongation factors TFIIF and ELL (Selby and Sancar, 1997). On the basis of these data they have suggested a model for the role of CSB. They suggest that through stimulation of transcript elongation, CSB is able to speed up the rate at which RNAP moves and therefore it may encounter any lesions present sooner than in the absence of CSB. This will subsequently lead to an increased rate of repair. Furthermore, CSB may also be able to position and stabilise RNAPII at a dimer and hence counteract the TFIIS-induced transcript degradation, (although no effect on TCR has been observed in yeast TFIIS mutants, (Verhage et al., 1997)).

This might go some way to explaining the faster rate of repair in the transcribed strand, but also suggests that CSB is involved in transcript elongation. As such, a slower rate of all RNAPII transcription might be expected in CS-B cells. Reduced rates of transcription have indeed been observed in CS cells (Balajee et al., 1997; Dianov et al., 1997), but this does not take into account other characteristics of CS cells that may indirectly affect the rate of transcription. CS cells have a shorter life span (Thompson and Holliday, 1983) and in many cases a slower growth rate (Arlett, Harcourt and Lehmann, unpublished observations). Therefore to obtain a true comparison of transcription rates, the number of cell divisions occurring over the course of the experiment must be taken into account. Furthermore, it has been suggested that TCR could allow growth to resume rapidly after genotoxic treatment in yeast (van Gool et al., 1994) which may result in an increase in the efficiency of RNAPII displacement and subsequently repair. Therefore CSB deficient cells

might have a slower recycling of RNAPII, which could result in a reduction in transcription rate.

7.4.1 A model for the role of CSB

It has been shown that approximately 10-15% of RNAPII binds tightly to CSB (Citterio et al., 1998) and this may be the proportion that is either stalled at a lesion or at a pause site. CSB may therefore preferentially bind to a stalled RNA polymerase molecule, thereby pinpointing the site of a lesion. Additionally, the binding of CSB to RNAPII may be a signal for ubiquitination of the large subunit of the C-terminal domain (Bregman et al., 1996), since CS cells are deficient in the damage induced ubiquitination of the large subunit of RNAPII. This will result in the degradation of the polymerase, and the subsequent removal of the stalled RNAP molecule, as in *E.coli*.

As discussed above, CSB may be able to translocate along the DNA using the energy derived from its ATPase activity and translocation may enable the protein to locate the site of a lesion or stalled polymerase (Tantin et al., 1997). This translocation activity and a possible ability to dissociate bound proteins could allow CSB to act as an uncoupling factor, as suggested for Rad26. CSB could move along the newly repaired DNA and dissociate the NER factors, allowing the resumption of transcription freeing TFIIH for participation in initiation.

E.coli RNAP, on encountering a pause site, is able to backtrack reversibly in an attempt to bypass it (Nudler et al., 1997). This leads to the possibility that mammalian RNAPII could backtrack from a lesion site, possibly helped by the ATPase activity of CSB, allowing the repair enzymes access to the site. However, the movement of RNAPII may not be required, since a stalled RNAPII molecule does not inhibit repair (Selby and Sancar, 1997). The assay, however was carried out using partially purified repair factors on an artificial substrate, and is not therefore able to detect a small retraction of RNAPII.

7.5 Concluding remarks

There are a considerable number of theories (discussed above) as to the function of CSB in both repair and transcription. This study has attempted to address some of these questions through a functional analysis of the gene. A number of regions, particularly the helicase/ATPase, glycine and acidic domains have been shown to be necessary for function. Further questions have also been raised, such as how the disruption of these regions results in loss of functional protein and whether any are specifically required for either repair of UV damage or enzymatic activity. It is hoped that continuation of this work will address these and other points.

REFERENCES

- Aboussekhra, A., M. Biggerstaff, M. K. K. Shivji, J. A. Vilpo, V. Moncollin, V. N. Podust, M. Protic, U. Hubscher, J.-M. Egly, and R. D. Wood. 1995. Mammalian DNA nucleotide excision repair reconstituted with purified components. *Cell* 80: 859-868.
- Adamczewski, J. P., M. Rossignol, J.-P. Tassan, E. A. Nigg, V. Moncollin, and J. M. Egly. 1996. MAT1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH. *The EMBO Journal* 15: 1877-1884.
- Akoulitchiev, S., T. P. Makela, R. A. Weinberg, and D. Reinberg. 1995. Requirement for TFIIH kinase activity in transcription by RNA polymerase II. *Nature* 377: 557-560.
- Arlett, C. F., and S. A. Harcourt. (1983). Variation in response to mutagens amongst normal and repair- defective human cells. In *Induced Mutagenesis. Molecular Mechanisms and their Implications for Environmental Protection*, C. W. Lawrence, ed. (New York: Plenum Press), pp. 249-266.
- Arlett, C. F., and A. R. Lehmann. 1979. The UV-response in repair-deficient human syndromes. *Acta. Biol. Med. Germ.* 338: 1277-1283.
- Arlett, C. F., A. R. Lehmann, F. Giannelli, and C. A. Ramsay. 1978. A human subject with a new defect in repair of ultraviolet damage. *J. Invest. Derm.* 70: 173-177.
- Asahina, H., I. Kuraoka, M. Shirakawa, E. H. Morita, N. Miura, I. Miyamoto, E. Ohtsuka, Y. Okada, and K. Tanaka. 1994. The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage. *Mutation Res.* 315: 229-237.
- Aso, T., J. W. Conaway, and R. C. Conaway. 1995. The RNA polymerase II elongation complex. *FASEB J.* 9: 1419-1428.
- Balajee, A. S., A. May, G. L. Dianov, E. C. Friedberg, and V. A. Bohr. 1997. Reduced RNA polymerase II transcription in intact and permeabilized Cockayne Syndrome group B cells. *Proc. Natl. Acad. Sci. USA* 94: 4306-4311.
- Bang, D., R. Verhage, N. Goosen, J. Brouwer, and P. van de Putte. 1992. Molecular cloning of RAD16, a gene involved in differential repair in *Saccharomyces cerevisiae*. *Nuc. Acids. Res.* 20: 3925-3931.
- Barnes, D. E., A. E. Tomkinson, A. R. Lehmann, A. D. B. Webster, and T. Lindahl. 1992. Mutations in the DNA ligase 1 gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA damaging agents. *Cell* 69: 495-503.
- Beal, P. A., and P. B. Dervan. 1991. Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. *Science* 251: 1360-1363.
- Bedoyan, J., R. Gupta, F. Thoma, and M. J. Smerdon. 1992. Transcription, nucleosome stability, and DNA repair in a yeast minichromosome. *J. Biol. Chem.* 267: 5996-6005.

- Bessho, T., A. Sancar, L. H. Thompson, and M. P. Thelen. 1997. Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex. *J. Biol. Chem.* 272: 3833-3837.
- Bhatia, P. K., R. A. Verhage, J. Brouwer, and E. C. Friedberg. 1996. Molecular cloning and characterization of *Saccharomyces cerevisiae* RAD28, the yeast homolog of the human Cockayne Syndrome A (CSA) gene. *J. Bacteriol.* 178: 5977-5988.
- Bockrath, R., and J. Palmer. 1977. Differential repair of premutational UV-lesions at *tRNA* genes in *E. coli*. *Molec. Gen. Genet.* 156: 133-140.
- Bohr, V. A., C. A. Smith, D. S. Okumoto, and P. C. Hanawalt. 1985. DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell* 40: 359-369.
- Bootsma, D., M. P. Mulder, E. Pot, and J. A. Cohen. 1970. Different inherited levels of DNA repair replication in xeroderma pigmentosum cell strains after exposure to ultraviolet irradiation. *Mutation Res.* 9: 507-516.
- Bregman, D. B., R. Halaban, A. J. van Gool, K. A. Henning, E. C. Friedberg, and S. L. Warren. 1996. UV-induced ubiquitination of RNA polymerase II: a novel modification deficient in Cockayne syndrome cells. *Proc. Natl. Acad. Sci. USA* 93: 11586-11590.
- Broughton, B. C., H. Steingrimsdottir, C. Weber, and A. R. Lehmann. 1994. Mutations in the xeroderma pigmentosum group D DNA repair gene in patients with trichothiodystrophy. *Nature Genetics* 7: 189-194.
- Broughton, B. C., A. F. Thompson, S. A. Harcourt, W. Vermeulen, J. H. J. Hoeijmakers, E. Botta, M. Stefanini, M. King, C. Weber, J. Cole, C. F. Arlett, and A. R. Lehmann. 1995. Molecular and cellular analysis of the DNA repair defect in a patient in xeroderma pigmentosum complementation group D with the clinical features of xeroderma pigmentosum and Cockayne Syndrome. *Am. J. Hum. Genet.* 56: 167-174.
- Buratowski, S. 1994. The basics of basal transcription by RNA polymerase II. *Cell* 77: 1-3.
- Burns, J. L., S. N. Guzder, P. Sung, S. Prakash, and L. Prakash. 1996. An affinity of human replication protein A for ultraviolet-damaged DNA. *J. Biol. Chem.* 271: 11607-11610.
- Cadena, L. D., and M. E. Dahmus. 1987. Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *J. Biol. Chem.* 262: 12468-12474.
- Cairns, B. R., Y. J. Kim, M. H. Sayre, B. C. Laurent, and R. D. Kornberg. 1994. A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5 and SNF6 gene products isolated from yeast. *Proc. Natl. Acad. Sci. USA* 91: 1950-1954.
- Cairns, B. R., Y. Lorch, Y. Li, M. Zhang, L. Lacomis, H. Erdjument-Bromage, P. Tempst, J. Du, B. Laurent, and R. D. Kornberg. 1996. RSC, an essential, abundant chromatin-remodeling complex. *Cell* 87: 1249-1260.

- Celis, J. E., and P. Madsen. 1986. Increased nuclear cyclin/PCNA antigen staining of non S-phase transformed human amnion cells engaged in nucleotide excision repair. *FEBS Lett.* **209**: 277-283.
- Chao, D. M., E. L. Gadbois, P. J. Murray, S. F. Anderson, M. S. Sonu, J. D. Parvin, and R. A. Young. 1996. A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature* **380**: 82-85.
- Christians, F. C., and P. C. Hanawalt. 1994. Repair in ribosomal RNA genes is deficient in xeroderma pigmentosum group C and in Cockayne's syndrome cells. *Mutation Res.* **323**: 179-187.
- Chu, G., and E. Chang. 1988. Xeroderma pigmentosum group E cells lack a Nuclear Factor that binds to damaged DNA. *Science* **242**: 564-567.
- Citterio, E., S. Rademakers, G. T. van der Horst, A. J. van Gool, J. H. Hoeijmakers, and W. Vermeulen. 1998. Biochemical and biological characterization of wild-type and ATPase-deficient Cockayne syndrome B repair protein. *J. Biol. Chem.* **273**: 11844-11851.
- Cleaver, J. E. 1968. Deficiency in repair replication of DNA in xeroderma pigmentosum. *Nature* **218**: 652-656.
- Cleaver, J. E. 1972. Xeroderma pigmentosum: variants with normal DNA repair and normal sensitivity to uv light. *J. Invest. Dermatol.* **58**: 124-128.
- Cleaver, J. E., and K. H. Kraemer. (1989). Xeroderma pigmentosum. In *The metabolic basis of inherited disease*, C. R. Scriver, A. L. Beaudet, W. S. Sly and D. Valle, eds. (New York: McGraw-Hill Inc), pp. 2949-2971.
- Cooper, P. K., T. Nospikel, S. G. Clarkson, and S. A. Leadon. 1997. Defective transcription-coupled repair of oxidative base damage in Cockayne syndrome patients from XP group G. *Science* **275**: 990-993.
- Coverley, D., M. K. Kenny, M. Munn, W. D. Rupp, D. P. Lane, and R. D. Wood. 1991. Requirement for the replication protein SSB in human DNA excision repair. *Nature* **349**: 538-541.
- Davis, J. L., R. Kunisawa, and J. Thorner. 1992. A presumptive helicase (MOT1 gene product) affects gene expression and is required for viability in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**: 1879-1892.
- de Laat, W. L., E. Appeldoorn, N. G. J. Jaspers, and J. H. J. Hoeijmakers. 1998. DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.* **273**: 7835-7842.
- de Vries, A., C. T. van Oostrom, F. M. Hofhuis, P. M. Dortant, R. J. Berg, F. R. de Gruijl, P. W. Wester, C. F. van Kreijl, P. J. Capel, H. van Steeg, and S. J. Verbeek. 1995. Increased susceptibility to ultraviolet-B and carcinogens of mice lacking the DNA excision repair gene XPA. *Nature* **377**: 169-173.
- Dianov, G. L., J. F. Houle, N. Iyer, V. A. Bohr, and E. C. Friedberg. 1997. Reduced RNA polymerase II transcription in extracts of Cockayne syndrome and xeroderma pigmentosum/Cockayne syndrome cells. *Nucl. Acids Res.* **25**: 3636-3642.

- Donahue, B. A., R. P. Fuchs, D. Reines, and P. C. Hanawalt. 1996. Effects of aminofluorene and acetylaminofluorene DNA adducts on transcriptional elongation by RNA polymerase II. *J. Biol. Chem.* **271**: 10588-10594.
- Donahue, B. A., S. Yin, J. S. Taylor, D. Reines, and P. C. Hanawalt. 1994. Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc. Natl. Acad. Sci. USA* **91**: 8502-8506.
- Drapkin, R., J. T. Reardon, A. Ansari, J. C. Huang, L. Zawel, K. Ahn, A. Sancar, and D. Reinberg. 1994. Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* **368**: 769-772.
- Drapkin, R., J. T. Reardon, A. Ansari, J.-C. Huang, L. Zawel, K. Ahn, A. Sancar, and D. Reinberg. 1994. Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* **368**: 769-772.
- Emery, H. S., D. Schild, D. E. Kellogg, and R. K. Mortimer. 1991. Sequence of RAD54, a *Saccharomyces cerevisiae* gene involved in recombination and repair. *Gene* **104**: 103-106.
- Evans, E., J. Fellows, A. Coffey, and R. D. Wood. 1997. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J.* **16**: 625-638.
- Evans, E., J. G. Moggs, J. R. Hwang, J. M. Egly, and R. D. Wood. 1997. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J.* **16**: 6559-6573.
- Evans, M. K., J. H. Robbins, M. B. Ganges, R. E. Tarone, R. S. Nairn, and B. V. A. 1993. Gene-specific DNA repair in xeroderma pigmentosum complementation groups A, C, D and F. Relation to cellular survival and clinical features. *J. Biol. Chem.* **268**: 4839-4847.
- Fairman, M. P., and B. Stillman. 1988. Cellular factors required for multiple stages of SV40 DNA replication *in vitro*. *EMBO J.* **7**: 1211-1218.
- Frederick, G. D., R. H. Amirkhan, R. A. Schultz, and E. C. Friedberg. 1994. Structural and mutational analysis of the xeroderma pigmentosum group D (XPD) gene. *Hum. Mol. Genet.* **3**: 1783-1788.
- Friedberg, E. C., G. C. Walker, and W. Siede. 1995. DNA repair and Mutagenesis. *ASM Press, Washington, USA*.
- Girdham, C. H., and D. M. Glover. 1991. Chromosome tangling and breakage at anaphase result from mutations in lodestar, a *Drosophila* gene encoding a putative nucleoside triphosphate-binding protein. *Genes Dev.* **5**: 1786-1799.
- Guzder, S. N., Y. Habraken, P. Sung, L. Prakash, and S. Prakash. 1996. RAD26, the yeast homolog of human Cockayne's syndrome group B gene, encodes a DNA-dependent ATPase. *J. Biol. Chem.* **271**: 18314-18317.
- Guzder, S. N., Y. Habraken, P. Sung, L. Prakash, and S. Prakash. 1995. Reconstitution of yeast nucleotide excision repair with purified Rad proteins, replication protein A, and transcription factor TFIIH. *J. Biol. Chem.* **270**: 12973-12976.

- Guzder, S. N., P. Sung, L. Prakash, and S. Prakash. 1993. Yeast DNA-repair gene *RAD14* encodes a zinc metalloprotein with affinity for ultraviolet-damaged DNA. *Proc. Natl. Acad. Sci. USA* 90: 5433-5437.
- Habraken, Y., P. Sung, L. Prakash, and S. Prakash. 1995. Structure-specific nuclease activity in yeast nucleotide excision-repair protein Rad2. *J. Biol. Chem.* 270: 30194-30198.
- Hanawalt, P. C. 1994. Transcription-coupled repair and human disease. *Science* 266: 1957-1958.
- Hanawalt, P. C., B. A. Donahue, and K. S. Sweder. 1994. Collision or collusion? *Curr. Biol.* 4: 518-521.
- He, Z., L. A. Henrickson, M. S. Wold, and C. J. Ingles. 1995. RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature* 374: 566-569.
- Henning, K. A., L. Li, N. Iyer, L. D. McDaniel, M. S. Reagan, R. Legerski, R. A. Schultz, M. Stefanini, A. R. Lehmann, L. V. Mayne, and E. C. Friedberg. 1995. The Cockayne-syndrome group-A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA-polymerase-II TFIIH. *Cell* 82: 555-564.
- Hirchhorn, J. N., S. A. Brown, C. D. Clark, and F. Winston. 1992. Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes Dev.* 6: 2288-2298.
- Hoeijmakers, J. H. J. 1993. Nucleotide excision repair 1: From Escherichia coli to yeast. *Trends Genet.* 9: 173-177.
- Hoeijmakers, J. H. J., J.-M. Egly, and W. Vermeulen. 1996. TFIIH: a key component in multiple DNA transactions. *Curr. Op. Gen. Dev.* 6: 26-33.
- Itoh, T., J. E. Cleaver, and M. Yamaizumi. 1996. Cockayne syndrome complementation group B associated with xeroderma pigmentosum phenotype. *Hum. Genet.* 97: 176-179.
- Itoh, T., T. Shiomi, N. Shiomi, Y. Harada, M. Wakasugi, T. Matsunaga, O. Nikaido, E. C. Friedberg, and M. Yamaizumi. 1996. Rodent complementation group 8 (ERCC8) corresponds to Cockayne syndrome complementation group A. *Mutation Res.* 362: 167-174.
- Iyer, N., M. S. Reagan, K. J. Wu, B. Canagarajah, and E. C. Friedberg. 1996. Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochem.* 35: 2157-2167.
- Jones, C. J., and R. D. Wood. 1993. Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochem.* 32: 12096-12104.
- Kassavetis, G. A., and E. P. Geiduschek. 1993. RNA polymerase marching backward. *Science* 259: 944-945.
- Keeney, S., A. P. M. Eker, T. Brody, W. Vermeulen, D. Bootsma, J. H. J. Hoeijmakers, and S. Linn. 1994. Correction of the DNA repair defect in xeroderma pigmentosum group

E by injection of a DNA damage-binding protein. *Proc. Natl. Acad. Sci. USA* 91: 4053-4056.

Keeney, S., H. Wein, and S. Linn. 1992. Biochemical heterogeneity in xeroderma pigmentosum complementation group E. *Mutation Res.* 273: 49-56.

Kim, C., R. O. Snyder, and M. S. Wold. 1992. Binding properties of replication protein A from human and yeast cells. *Mol. Cell. Biol.* 12: 3050-3059.

Koleske, A. J., and R. A. Young. 1995. The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci.* 20: 113-116.

Koleske, A. J., and R. A. Young. 1994. An RNA polymerase II holoenzyme responsive to activators. *Nature* 368: 466-469.

Kraemer, K. H., and H. Slor. 1984. Xeroderma pigmentosum. *Clinics in Dermatol.* 2: 33-69.

Krude, T., M. Jackman, J. Pines, and R. A. Laskey. 1997. Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. *Cell* 88: 109-119.

Kruger, W., C. L. Peterson, A. Sil, C. Coburn, G. Arents, E. N. Moudrianakis, and I. Herskowitz. 1995. Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. *Genes Dev.* 9: 2770-2779.

Kunala, S., and D. E. Brash. 1995. Interagenic domains of strand-specific repair in *Escherichia coli*. *J. Mol. Biol.* 246: 264-272.

Kuraoka, I., E. H. Morita, M. Saijo, T. Matsuda, K. Morikawa, M. Shirakawa, and K. Tanaka. 1996. Identification of a damaged-DNA binding domain of the XPA protein. *Mutation Res.* 362: 87-95.

Kwon, H., A. N. Imbalzano, P. A. Khavari, R. E. Kingston, and M. R. Green. 1994. Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. *Nature* 370: 477-481.

Laurent, B. C., I. Treich, and M. Carlson. 1993. The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. *Genes Dev.* 7: 583-591.

Leadon, S. A., and A. V. Avrutskaya. 1997. Differential involvement of the human mismatch repair proteins, hMLH1 and hMSH2, in transcription-coupled repair. *Cancer Res.* 57: 3784-3791.

Leadon, S. A., and A. V. Avrutskaya. 1998. Requirement for DNA mismatch repair proteins in the transcription-coupled repair of thymine glycols in *Saccharomyces cerevisiae*. *Mutation Res.* 407: 177-187.

Leadon, S. A., and P. K. Cooper. 1993. Preferential repair of ionizing radiation-induced damage in the transcribed strand of an active human gene is defective in Cockayne syndrome. *Proc. Natl. Acad. Sci. USA* 90: 10499-10503.

- Leadon, S. A., and D. A. Lawrence. 1991. Preferential repair of DNA damage on the transcribed strand of the human metallothionein genes requires RNA polymerase II. *Mutation Res.* 255: 67-78.
- Leadon, S. A., and D. A. Lawrence. 1992. Strand-selective repair of DNA damage in the yeast GAL7 gene requires RNA polymerase II. *J. Biol. Chem.* 267: 23175-23182.
- Lehmann, A. R. 1995. Nucleotide excision repair and the link with transcription. *Trends Biochem. Sci.* 20: 402-405.
- Lehmann, A. R. 1982. Three complementation groups in Cockayne syndrome. *Mutation Res.* 106: 347-356.
- Lehmann, A. R., S. Kirk-Bell, C. F. Arlett, M. C. Paterson, P. H. M. Lohman, E. A. de Weerd-Kastelein, and D. Bootsma. 1975. Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. *Proc. Natl. Acad. Sci. USA* 72: 219-223.
- Li, B.-H., and R. Bockrath. 1995. Mutation frequency decline in *Escherichia coli*. *Molec. Gen. Genet.* 249: 585-599.
- Li, L., E. S. Bales, C. A. Peterson, and R. J. Legerski. 1993. Characterization of molecular defects in xeroderma pigmentosum group C. *Nature Genetics* 5: 413-417.
- Li, L., X. Lu, C. A. Peterson, and R. J. Legerski. 1995. An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Mol. Cell. Biol.* 15: 5396-5402.
- Lieber, M. R. 1997. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* 19: 233-240.
- Lowry, R. B. 1982. Early onset of Cockayne syndrome. *American Journal of Medical Genetics* 13: 209-10.
- Lu, H., L. Zawel, L. Fisher, J.-M. Egly, and D. Reinberg. 1992. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358: 641-645.
- Ma, L., A. Westbroek, A. G. Jochemsen, G. Weeda, A. Bosch, D. Bootsma, J. H. J. Hoeijmakers, and A. J. van der Eb. 1994. Mutational analysis of ERCC3, which is involved in DNA repair and transcription initiation: identification of domains essential for the DNA repair function. *Mol. Cell. Biol.* 14: 4126-4134.
- Madhani, H. D., V. A. Bohr, and P. C. Hanawalt. 1986. Differential DNA repair in transcriptionally active and inactive proto-oncogenes: c-abl and c-mos. *Cell* 45: 417-423.
- Makela, T. P., J.-P. Tassan, E. A. Nigg, S. Frutiger, G. J. Hughes, and R. A. Weinberg. 1994. A cyclin associated with the CDK-activating kinase MO15. *Nature* 371: 254-257.
- Maldonado, E., R. Shiekhatar, M. Sheldon, H. Cho, R. Drapkin, P. Rickert, E. Lees, C. W. Anderson, S. Linn, and D. Reinberg. 1996. A human RNA polymerase II complex associated with SRB and DNA-repair proteins. *Nature* 381: 86-89.

- Mallery, D. L., B. Tanganelli, S. Colella, H. Steingrimsdottir, A. J. Van Gool, C. Troelstra, M. Stefanini, and A. R. Lehmann. 1998. Molecular analysis of mutations in the *CSB(ERCC6)* gene in patients with Cockayne Syndrome. *Am. J. Hum. Genet.* 62: 77-85.
- Marinoni, J.-C., R. Roy, W. Vermeulen, P. Miniou, Y. Lutz, G. Weeda, T. Seroz, D. M. Gomez, J. H. J. Hoeijmakers, and J.-M. Egly. 1997. Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH. *EMBO J.* 16.
- Masutani, C., K. Sugasawa, J. Yanagisawa, T. Sonoyama, M. Ui, T. Enomoto, K. Takio, K. Tanaka, P. J. van der Spek, D. Bootsma, J. H. J. Hoeijmakers, and F. Hanaoka. 1994. Purification and cloning of a nucleotide excision-repair complex involving the xeroderma-pigmentosum group-C protein and a human homolog of yeast RAD23. *EMBO J.* 13: 1831-1843.
- Matsunaga, T., D. Mu, C.-H. Park, J. T. Reardon, and A. Sancar. 1995. Human DNA-repair excision nuclease. *J. Biol. Chem.* 270: 20862-20869.
- Matsunaga, T., C.-H. Park, T. Bessho, D. Mu, and A. Sancar. 1996. Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. *J. Biol. Chem.* 271: 11047-11050.
- Mayne, L. V., and A. R. Lehmann. 1982. Failure of RNA Synthesis to Recover after UV-irradiation: An Early Defect in Cells from Individuals with Cockayne's Syndrome and Xeroderma Pigmentosum. *Cancer Res.* 42: 1473-1478.
- Mayne, L. V., A. R. Lehmann, and R. Waters. 1982. Excision repair in Cockayne syndrome. *Mutation Res.* 106: 179-189.
- McGregor, W. G., C.-M. Mah, R.-H. Chen, V. M. Mayer, and J. J. McCormick. 1995. Lack of correlation between degree of interference with transcription and rate of strand specific repair in the HPRT gene of diploid human fibroblasts. *J. Biol. Chem.* 270: 27222-27227.
- McWhir, J., J. Selfridge, D. J. Harrison, S. Squires, and D. W. Melton. 1993. Mice with DNA repair gene (*ERCC-1*) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nature Genetics* 5: 217-223.
- Mellon, I., and G. N. Champe. 1996. Products of DNA mismatch repair genes *MutS* and *MutL* are required for transcription-coupled nucleotide-excision repair of the lactose operon in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 93: 1292-1297.
- Mellon, I., and P. C. Hanawalt. 1989. Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* 342: 95-98.
- Mellon, I., D. K. Rajpal, M. Koi, C. R. Boland, and G. N. Champe. 1996. Transcription-coupled repair deficiency and mutations in human mismatch repair genes. *Science* 272: 557-560.
- Mellon, I., G. Spivak, and P. C. Hanawalt. 1987. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* 51: 241-249.

- Miyamoto, I., N. Miura, H. Niwa, J. Miyazaki, and K. Tanaka. 1992. Mutational analysis of the structure and function of the xeroderma pigmentosum group A complementing protein. *J. Biol. Chem.* **267**: 12182-12187.
- Miyauchi, H., T. Horio, T. Akaeda, Y. Asada, H. R. Chang, K. Ishizaki, and M. Ikenaga. 1994. Cockayne syndrome in two adult siblings. *J. Am. Acad. Dermatol.* **30**: 329-335.
- Mizuguchi, G., T. Tsukiyama, J. Wisniewski, and C. Wu. 1997. Role of nucleosome remodeling factor NURF in transcriptional activation of chromatin. *Molecular Cell* **1**: 141-150.
- Moggs, J. G., K. J. Yarema, J. M. Essigmann, and R. D. Wood. 1996. Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *J. Biol. Chem.* **271**: 7177-7186.
- Mu, D., D. S. Hsu, and A. Sancar. 1996. Reaction mechanism of human DNA repair excision nuclease. *J. Biol. Chem.* **271**: 8285-8294.
- Mu, D., C. H. Park, T. Matsunaga, D. S. Hsu, J. T. Reardon, and A. Sancar. 1995. Reconstitution of human DNA-repair excision nuclease on a highly defined system. *J. Biol. Chem.* **270**: 2415-2418.
- Mu, D., and A. Sancar. 1997. Model for XPC-independent transcription-coupled repair of pyrimidine dimers in humans. *J. Biol. Chem.* **272**: 7570-7573.
- Nagai, A., M. Saijo, I. Kuraoka, T. Matsuda, N. Kodo, Y. Nakatsu, T. Mimaki, M. Mino, M. Biggerstaff, R. D. Wood, A. Sijbers, J. H. J. Hoeijmakers, and K. Tanaka. 1995. Enhancement of damage-specific DNA binding of XPA by interaction with the ERCC1 DNA repair protein. *Biochem. Biophys. Res. Commun.* **211**: 960-966.
- Nakane, H., S. Takeuchi, S. Yuba, M. Sijo, Y. Nakatsu, H. Murai, Y. Nakatsuru, T. Ishikawa, S. Hirota, Y. Kitamura, Y. Kato, Y. Tsunoda, H. Miyauchi, T. Horio, T. Tokunaga, T. Matsunaga, O. Nikaido, Y. Nishimune, Y. Okada, and K. Tanaka. 1995. High incidence of ultraviolet-B- or chemical-carcinogen- induced skin tumours in mice lacking the xeroderma pigmentosum group A gene. *Nature* **377**: 165-168.
- Nance, M. A., and S. A. Berry. 1992. Cockayne Syndrome: Review of 140 cases. *Am. J. Med. Genet.* **42**: 68-84.
- Neer, E. J., C. J. Schmidt, R. Nambudripad, and T. F. Smith. 1994. The ancient regulatory-protein family of WD-repeat proteins. *Nature* **371**: 297-300.
- Nichols, A. F., P. Ong, and S. Linn. 1996. Mutations specific to the xeroderma pigmentosum group E Ddb- phenotype. *J. Biol. Chem.* **40**: 24317-24320.
- Nichols, A. F., and A. Sancar. 1992. Purification of PCNA as a nucleotide excision repair protein. *Nucl. Acids Res.* **20**: 3559-3564.
- Nikolov, D. B., and S. K. Burley. 1997. RNA polymerase II transcription initiation: A structural view. *Proc. Natl. Acad. Sci. USA* **94**: 15-22.

- Nudler, E., A. Mustaev, E. Lukhtanov, and A. Goldfarb. 1997. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* 89: 33-41.
- O'Brien, T., S. Hardin, A. Greenleaf, and J. T. Lis. 1994. Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature* 370: 75-77.
- O'Donovan, A., A. A. Davies, J. G. Moggs, S. C. West, and R. D. Wood. 1994. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* 371: 432-435.
- O'Donovan, A., and R. D. Wood. 1993. Identical defect in DNA repair in xeroderma pigmentosum group G and rodent ERCC group 5. *Nature* 363: 185-188.
- Okhuma, Y., and R. G. Roeder. 1994. Regulation of TFIIH ATPase and kinase activities by TFIIIE during active initiation of complex formation. *Nature* 368: 160-163.
- Ossipow, V., J. P. Tassan, E. A. Nigg, and U. Schibler. 1995. A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. *Cell* 83: 137-146.
- Owen-Hughes, T., R. T. Utley, J. Cote, C. L. Peterson, and J. L. Workman. 1996. Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. *Science* 273: 513-516.
- Park, C.-H., D. Mu, J. T. Reardon, and A. Sancar. 1995. The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIIE transcription factor. *J. Biol. Chem.* 270: 4896-4902.
- Park, C.-H., and A. Sancar. 1993. Reconstitution of mammalian excision repair ability with mutant cell-free extracts and XPAC and ERCC1 proteins expressed in *Escherichia coli*. *Nucl. Acids Res.* 21: 5110-5116.
- Patterson, M., and G. Chu. 1989. Evidence that XP cells from complementation group E are deficient in a homolog of yeast photolyase. *Mol. Cell. Biol.* 9: 5105-5112.
- Pavco, P. A., and D. A. Steege. 1990. Elongation by *Escherichia coli* RNA polymerase is blocked *in vitro* by a site-specific DNA binding protein. *Journal of Biological Science* 265: 9960-9969.
- Pazin, M. J., and J. T. Kadonaga. 1997. SWI2/SNF2 and related proteins: ATP-driven motors that disrupt protein-DNA interactions. *Cell* 88: 737-740.
- Peterson, C. L., A. Dingwall, and M. P. Scott. 1994. Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. *Proc. Natl. Acad. Sci. USA* 91: 2905-2908.
- Podust, L. M., V. N. Podust, C. Floth, and U. Hubscher. 1994. Assembly of DNA polymerase δ and ϵ holoenzymes depends on the geometry of the DNA template. *Nucl. Acids Res.* 22: 2970-2975.

- Pollitt, R. J., F. A. Fenner, and M. Davies. 1968. Sibs with mental and physical retardation and trichorrexis nodosa with abnormal amino acid composition of the hair. *Arch. Dis. Child.* 43: 211-216.
- Ramanathan, B., and M. J. Smerdon. 1989. Enhanced DNA repair synthesis in hyperacetylated nucleosomes. *J. Biol. Chem.* 264: 11026-11034.
- Ranish, J. A., and S. Hahn. 1996. Transcription: basal factors and activation. *Curr. Op. Gen. Dev.* 6: 151-158.
- Rathmell, W. K., and G. Chu. 1994. A DNA end-binding factor involved in double-strand break repair and V(D)J recombination. *Mol. Cell. Biol.* 14: 4741-4748.
- Ratner, J. N., B. Balasubramanian, J. Corden, S. L. Warren, and D. B. Bregman. 1998. Ultraviolet radiation-induced ubiquitination and proteasomal degradation of the large subunit of RNA polymerase II. Implications for transcription-coupled DNA repair. *J. Biol. Chem.* 273: 5184-5189.
- Reagan, M. S., and E. C. Friedberg. 1997. Recovery of RNA polymerase II synthesis following DNA damage in mutants of *Saccharomyces cerevisiae* defective in nucleotide excision repair. *Nucl. Acids Res.* 25: 4257-4263.
- Reardon, J., L. H. Thompson, and A. Sancar. 1993. Excision repair in man and the molecular basis of xeroderma pigmentosum syndrome. *Cold Spring Harbor Symp. Quant. Biol.* 58: 605-617.
- Reardon, J. T., D. Mu, and A. Sancar. 1996. Overproduction, purification and characterization of the XPC subunit of the human DNA repair excision nuclease. *J. Biol. Chem.* 271: 19451-19456.
- Reines, D., J. W. Conaway, and R. C. Conaway. 1996. The RNA polymerase II general elongation factors. *Trends Biochem. Sci.* 21: 351-355.
- Richmond, E., and C. L. Peterson. 1996. Functional analysis of the DNA-stimulated ATPase domain of yeast SWI2/SNF2. *Nucl. Acids Res.* 24: 3685-3692.
- Robins, P., C. J. Jones, M. Biggerstaff, T. Lindahl, and R. D. Wood. 1991. Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA. *EMBO J.* 10: 3913-3921.
- Rossignol, M., I. Kolb-Cheynel, and J.-M. Egly. 1997. Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH. *EMBO J.* 16: 1628-1637.
- Roy, R., J. P. Adamczewski, T. Seroz, W. Vermeulen, J.-P. Tassan, L. Schaeffer, E. A. Nigg, J. H. J. Hoeijmakers, and J.-M. Egly. 1994. The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. *Cell* 79: 1093-1101.
- Sambrook, J., E. F. Fritsch, and T. E. Maniatis. (1989). *Molecular Cloning: a laboratory manual* (2nd Edition): Cold Spring Harbor Laboratory Press).
- Satokata, I., K. Tanaka, N. Miura, I. Miyamoto, Y. Satoh, S. Kondo, and Y. Okada. 1990. Characterization of a splicing mutation in group A xeroderma pigmentosum. *Proc. Natl. Acad. Sci. USA* 87: 9908-9912.

- Satokata, I., K. Tanaka, N. Miura, M. Narita, T. Mimaki, Y. Satoh, S. Kondo, and Y. Okada. 1992. Three nonsense mutations responsible for group A xeroderma pigmentosum. *Mutation Res.* 273: 193-202.
- Satokata, I., K. Tanaka, S. Yuba, and Y. Okada. 1992. Identification of splicing mutations of the last nucleotides of exons, a nonsense mutation, and a missense mutation of the XPAC gene as causes of group A xeroderma pigmentosum. *Mutation Res.* 273: 203-212.
- Schaeffer, L., V. Moncollin, R. Roy, A. Staub, M. Mezzina, A. Sarasin, G. Weeda, J. H. J. Hoeijmakers, and J. M. Egly. 1994. The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.* 13: 2388-2392.
- Schaeffer, L., R. Roy, S. Humbert, V. Moncollin, W. Vermeulen, J. H. J. Hoeijmakers, P. Chambon, and J.-M. Egly. 1993. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* 260: 58-63.
- Schiestl, R. H., and S. Prakash. 1990. RAD10, an excision repair gene of *Saccharomyces cerevisiae*, is involved in the RAD1 pathway of mitotic recombination. *Mol. Cell. Biol.* 10: 2485-2491.
- Schild, D., B. J. Glassner, R. K. Mortimer, M. Carlson, and B. C. Laurents. 1992. Identification of RAD16, a yeast excision repair gene homologous to the recombinational repair gene RAD54 and to the SNF2 gene involved in transcription activation. *Yeast* 8: 385-395.
- Selby, C. P., and A. Sancar. 1997. Cockayne syndrome group B protein enhances elongation by RNA polymerase II. *Proc. Natl. Acad. Sci. USA* 95: 11205-11209.
- Selby, C. P., and A. Sancar. 1997. Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *J. Biol. Chem.* 272: 1885-1890.
- Selby, C. P., and A. Sancar. 1993. Molecular mechanism of transcription repair coupling. *Science* 260: 53-58.
- Selby, C. P., and A. Sancar. 1995. Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. *J. Biol. Chem.* 270: 4882-4889.
- Selby, C. P., and A. Sancar. 1995. Structure and function of transcription-repair coupling factor. II. Catalytic properties. *J. Biol. Chem.* 270: 4890-4895.
- Selby, C. P., and A. Sancar. 1990. Transcription preferentially inhibits nucleotide excision repair of the template DNA strand *in vitro*. *J. Biol. Chem.* 265: 21330-21336.
- Selby, C. P., E. M. Witkin, and A. Sancar. 1991. *Escherichia coli mfd* mutant deficient in "mutation frequency decline" lacks strand-specific repair: *In vitro* complementation with purified coupling factor. *Proc. Natl. Acad. Sci. USA* 88: 11574-11578.
- Serizawa, H., J. W. Conaway, and R. C. Conaway. 1994. An oligomeric form of the large subunit of transcription factor (TF)IIH activates phosphorylation of the RNA polymerase II carboxyl-terminal domain by TFIIH. *J. Biol. Chem.* 269: 20750-20756.

- Serizawa, H., T. M. Makela, J. W. Conaway, and R. C. Conaway. 1995. Association of cdk-activating kinase subunits with transcription factor TFIIH. *Nature* 374: 280-282.
- Shiekhataar, R., F. Mermelstein, R. P. Fisher, R. Drapkin, B. Dynlacht, H. C. Wesselin, D. O. Morgan, and D. Reinberg. 1995. Cdk-activating kinase complex is a component of human transcription factor TFIIH. *Nature* 374: 283-287.
- Shivji, M. K. K., M. K. Kenny, and R. D. Wood. 1992. Proliferating cell nuclear antigen is required for DNA excision repair. *Cell* 69: 367-374.
- Sijbers, A. M., W. L. de Laat, R. R. Ariza, M. Biggerstaff, Y.-F. Wei, J. G. Moggs, K. C. Carter, B. K. Shell, E. Evans, M. C. de Jong, S. Rademakers, J. de Rooij, N. G. J. Jaspers, J. H. J. Hoeijmakers, and R. D. Wood. 1996. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* 86: 811-822.
- Smerdon, M. J., and F. Thoma. 1990. Site-specific DNA repair at the nucleosome level in a yeast minichromosome. *Cell* 61: 675-684.
- St. John, T. P., S. Scherer, M. W. McDonell, and R. W. Davis. 1981. Deletion analysis of the *Saccharomyces* GAL gene cluster. Transcription from three promoters. *J. Mol. Biol.* 152: 317-334.
- Stargell, L. A., and K. Struhl. 1996. Mechanisms of transcriptional activation *in vivo*: two steps forward. *Trends Genet.* 12: 311-315.
- Stefanini, M., H. Fawcett, E. Botta, T. Nardo, and A. R. Lehmann. 1996. Genetic analysis of twenty-two patients with Cockayne syndrome. *Hum. Genet.* 97: 418-423.
- Stefanini, M., P. Lagomarsini, C. F. Arlett, S. Marinoni, C. Borrone, F. Crovato, G. Trevisan, G. Cordone, and F. Nuzzo. 1986. Xeroderma pigmentosum (complementation group D) mutation is present in patients affected by trichothiodystrophy with photosensitivity. *Hum. Genet.* 74: 107-112.
- Stefanini, M., P. Lagomarsini, R. Giorgi, and F. Nuzzo. 1987. Complementation studies in cells from patients affected by trichothiodystrophy with normal or enhanced photosensitivity. *Mutation Res.* 191: 117-119.
- Stefanini, M., W. Vermeulen, G. Weeda, S. Giliani, T. Nardo, M. Mezzina, A. Sarasin, J. I. Harper, C. F. Arlett, J. H. J. Hoeijmakers, and A. R. Lehmann. 1993. A new nucleotide excision repair gene associated with the genetic disorder trichothiodystrophy. *Am. J. Hum. Genet.* 53: 817-821.
- Sugasawa, K., C. Masutani, A. Uchida, T. Maekawa, P. J. van der Spek, D. Bootsma, J. H. J. Hoeijmakers, and F. Hanaoka. 1996. HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair *in vitro*. *Mol. Cell. Biol.* 16: 4852-4861.
- Sugasawa, K., J. M. Y. Ng, C. Masutani, T. Maekawa, A. Uchida, P. J. van der Spek, A. P. M. Eker, S. Rademakers, C. Visser, A. Aboussekhra, R. D. Wood, F. Hanaoka, D. Bootsma, and J. H. J. Hoeijmakers. 1997. Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity. *Mol. Cell. Biol.* 17: 6924-6931.

- Sweder, K. S., and P. C. Hanawalt. 1992. Preferential repair of cyclobutane pyrimidine dimers in the transcribed strand of a gene in yeast chromosomes and plasmids is dependent on transcription. *Proc. Natl. Acad. Sci. USA* 89: 10696-10700.
- Sweder, K. S., R. A. Verhage, D. J. Crowley, G. F. Crouse, J. Brouwer, and P. C. Hanawalt. 1996. Mismatch repair mutants in yeast are not defective in transcription-coupled DNA repair of UV-induced DNA damage. *Genetics* 143: 1127-1135.
- Takayama, K., E. P. Salazar, B. C. Broughton, A. R. Lehmann, A. Sarasin, L. H. Thompson, and C. A. Weber. 1996. Defects in the DNA repair and transcription gene *ERCC2(XPD)* in trichothiodystrophy. *Am. J. Hum. Genet.* 58: 263-270.
- Takayama, K., E. P. Salazar, A. R. Lehmann, M. Stefanini, L. H. Thompson, and C. A. Weber. 1995. Defects in the repair and transcription gene *ERCC2* in the cancer-prone disorder xeroderma pigmentosum. *Cancer Res.* 55: 5656-5663.
- Tanaka, K., K. Kawai, Y. Kumahara, and M. Ikenaga. 1981. Genetic Complementation Groups in Cockayne Syndrome. *Somatic Cell Genet.* 7: 445-456.
- Tanaka, K., M. Naoyuki, I. Satokata, I. Miyamoto, M. C. Yoshida, Y. Satoh, S. Kondo, A. Yasui, H. Okayama, and Y. Okada. 1990. Analysis of a human DNA excision repair gene involved in group A xeroderma pigmentosum and containing a zinc-finger domain. *Nature* 348: 73-76.
- Tansey, W. P., and W. Herr. 1997. TAFs: Guilt by association? *Cell* 88: 729-732.
- Tantin, D., A. Kansal, and M. Carey. 1997. Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Mol. Cell. Biol.* 17: 6803-6814.
- Taylor, E., B. C. Broughton, E. Botta, M. Stefanini, A. Sarasin, N. G. J. Jaspers, H. Fawcett, S. A. Harcourt, C. F. Arlett, and A. R. Lehmann. 1997. Xeroderma pigmentosum and trichothiodystrophy are associated with different mutations in the *XPD (ERCC2)* repair/transcription gene. *Proc. Natl. Acad. Sci. USA* 94: 8658-8663.
- Terleth, C., C. A. van Sluis, and P. van de Putte. 1989. Differential repair of UV damage in *Saccharomyces cerevisiae*. *Nuc. Acids. Res.* 17: 4433-4439.
- Thompson, K. V. A., and R. Holliday. 1983. Genetic Effects on the Longevity of Cultured Human Fibroblasts. II. DNA Repair Deficient Syndromes. *Gerontology* 29: 83-88.
- Tijsterman, M., R. J. Verhage, P. van de Putte, J. G. Tasseron-de Jong, and J. Brouwer. 1997. Transitions in the coupling of transcription and nucleotide excision repair within RNA polymerase II-transcribed genes of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 94: 8027-8032.
- Tjian, R., and T. Maniatis. 1994. Transcriptional activation: a complex puzzle with few easy pieces. *Cell* 77: 5-8.
- Troelstra, C., H. Odijk, J. de Wit, A. Westerveld, L. H. Thompson, D. Bootsma, and J. H. J. Hoeijmakers. 1990. Molecular cloning of the human DNA repair gene *ERCC6*. *Mol. Cell. Biol.* 10: 5806-5813.

- Troelstra, C., A. van Gool, J. de Wit, W. Vermeulen, D. Bootsma, and J. H. J. Hoeijmakers. 1992. ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. *Cell* 71: 939-953.
- van der Horst, G. T. J., H. van Steeg, R. J. W. Berg, A. J. van Gool, J. de Wit, G. Weeda, H. Morreau, R. B. Beems, C. F. van Kreijl, F. R. de Gruijl, D. Bootsma, and J. H. J. Hoeijmakers. 1997. Defective transcription-coupled repair in Cockayne syndrome B mice is associated with skin cancer predisposition. *Cell* 89: 425-435.
- van Gool, A. J., E. Citterio, S. Rademakers, R. van Os, W. Vermeulen, A. Constantinou, J.-M. Egly, D. Bootsma, and H. J. Hoeijmakers. 1997. The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in a RNA polymerase II containing complex. *EMBO J.* 16: 5955-5965.
- van Gool, A. J., G. T. J. van der Horst, E. Citterio, and J. H. J. Hoeijmakers. 1997. Cockayne syndrome: defective repair of transcription. *EMBO J.* 16: 4155-4162.
- van Gool, A. J., R. Verhage, S. M. A. Swagemakers, P. van de Putte, J. Brouwer, C. Troelstra, D. Bootsma, and J. H. J. Hoeijmakers. 1994. RAD26, the functional *S. cerevisiae* homolog of the cockayne syndrome B gene *ERCC6*. *EMBO J.* 13: 5361-5369.
- van Oosterwijk, M. F., A. Versteeg, R. Filon, A. A. van Zeeland, and L. H. F. Mullenders. 1996. The sensitivity of Cockayne's Syndrome cells to DNA-damaging agents is not due to defective transcription-coupled repair of active genes. *Mol. Cell. Biol.* 16: 4436-4444.
- van Vuuren, A. J., W. Vermeulen, L. Ma, G. Weeda, E. Appeldoorn, N. G. J. Jaspers, v. d. E. A. J., D. Bootsma, J. H. J. Hoeijmakers, S. Humbert, L. Schaeffer, and J. M. Egly. 1994. Correction of xeroderma pigmentosum repair defect by basal transcription factor BTF2 (TFIIH). *EMBO J.* 13: 1645-1653.
- Venema, J., Z. Bartosova, A. T. Natarajan, A. A. van Zeeland, and L. H. F. Mullenders. 1992. Transcription affects the rate but not the extent of repair of cyclobutane pyrimidine dimers in the human adenosine deaminase gene. *J. Biol. Chem.* 267: 8852-8856.
- Venema, J., L. H. F. Mullenders, A. T. Natarajan, A. A. Van Zeeland, and L. V. Mayne. 1990. The genetic defect in Cockayne's syndrome is associated with a defect in repair of uv-induced DNA damage in transcriptionally active DNA. *Proc. Natl. Acad. Sci. USA* 87: 4707-4711.
- Venema, J., A. van Hoffen, V. Karcagi, A. T. Natarajan, A. A. van Zeeland, and L. H. F. Mullenders. 1991. Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. *Mol. Cell. Biol.* 11: 4128-4134.
- Venema, J. A., A. T. Van Hoffen, A. A. Natarajan, V. Zeeland, and L. Mullenders. 1990. The residual repair capacity of xeroderma pigmentosum complementation group C fibroblasts is highly specific for transcriptionally active DNA. *Nucl. Acids Res.* 18: 443-448.
- Verhage, R., A. M. Zeeman, N. De Groot, F. Gleig, D. D. Bang, P. Van de Putte, and J. Brouwer. 1994. The *RAD7* and *RAD16* genes, which are essential for pyrimidine dimer removal from the silent mating type loci, are also required for repair of the nontranscribed strand of an active gene in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 14: 6135-6142.

- Verhage, R. A., J. Heyn, P. van de Putte, and J. Brouwer. 1997. Transcription elongation factor S-II is not required for transcription-coupled repair in yeast. *Molec. Gen. Genet.* **254**: 284-290.
- Verhage, R. A., A. J. van Gool, N. de Groot, J. H. J. Hoeijmakers, P. van de Putte, and J. Brouwer. 1996. Double mutants of *Saccharomyces cerevisiae* with alterations in global genome and transcription-coupled repair. *Mol. Cell. Biol.* **16**: 496-502.
- Verhage, R. A., A.-M. Zeeman, M. Lombaerts, P. van de Putte, and J. Brouwer. 1996. Analysis of gene- and strand-specific repair in the moderately UV sensitive *Saccharomyces cerevisiae* rad23 mutant. *Mutation Res.* **362**: 155-165.
- Vermeulen, W., J. Jaeken, N. G. J. Jaspers, D. Bootsma, and J. H. J. Hoeijmakers. 1993. Xeroderma pigmentosum complementation group G associated with Cockayne syndrome. *Am. J. Hum. Genet.* **53**: 185-192.
- Vermeulen, W., A. J. van Vuuren, M. Chipoulet, L. Schaeffer, E. Appeldoorn, G. Weeda, N. G. J. Jaspers, A. Priestley, C. F. Arlett, A. R. Lehmann, M. Stefanini, M. Mezzina, A. Sarasin, D. Bootsma, J.-M. Egly, and J. H. J. Hoeijmakers. 1994. Three unusual repair deficiencies associated with transcription factor BTF2(TFIIH): evidence for the existence of a transcription syndrome. *Cold Spring Harbor Symp. Quant. Biol.* **59**: 317-329.
- Wang, G., M. M. Seidman, and P. M. Glazer. 1996. Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science* **271**: 802-805.
- Weeda, G., I. Donker, J. de Wit, H. Morreau, R. Janssens, C. J. Vissers, A. Nigg, H. van Steeg, D. Bootsma, and J. H. J. Hoeijmakers. 1997. Disruption of mouse *ERCC1* results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr. Biol.* **7**: 427-439.
- Weeda, G., R. C. A. van Ham, W. Vermeulen, D. Bootsma, A. J. van der Eb, and J. H. J. Hoeijmakers. 1990. A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's Syndrome. *Cell* **62**: 777-791.
- Witkin, E. M. 1994. Mutation frequency decline revisited. *Bioessays* **16**: 437-444.
- Wold, M. S., and T. Kelly. 1988. Purification and characterization of replication protein A, a cellular protein required for *in vitro* replication of simian virus 40 DNA. *Proc. Natl. Acad. Sci. USA* **85**: 2523-2527.
- Yankulov, K. Y., and D. L. Bentley. 1997. Regulation of cdk7 substrate specificity by MAT1 and TFIIH. *EMBO J.* **16**: 1638-1646.
- You, Z., W. J. Feaver, and E. C. Friedberg. 1998. Yeast RNA polymerase II transcription *in vitro* is inhibited in the presence of nucleotide excision repair: complementation of inhibition by holo-TFIIH and requirement for RAD26. *Mol. Cell. Biol.* **18**: 2668-2676.
- Zawel, L., K. P. Kumar, and D. Reinberg. 1995. Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* **9**: 1479-1490.

APPENDIX I The primers used during this study

PRIMER	LOCATION	USE	SEQUENCE
DM10	2129-2150	exon 10 3' PCR primer	AGTCAAAGAGCGACCACAGCTC
DM11	n/a	exon 10 5' PCR primer	GGAAATGGCGATTGCTAAAAGGTAA
DM20	n/a	intron 11 PCR	TTTAAATATTAAATGCCAAAGTA
DM23	200-224	213 mutagenesis	AACGAAAGACGGGTACTCCTCCAC
DM24	2075-2099	2087 mutagenesis	ACAGAAATGATCCAATGAGGGGTGCG
DM25	4468-4494	4481 term mutagenesis	TTTCTCAATAGTTATCGGAAGACACAA
DM26	741-780	Sph site mutagenesis	GTCTCCTGGACAGGCATGAGCATGCTGCCAAGACTGGATG
DM27	58-97	ATG 80 mutagenesis	GGCCCTGGGTAGTCTGTAGAGAAATGCCAAATGAGGGAATC
DM28	145-184	ATG 164 mutagenesis	TGTCAGTAATAATGAAGAAATGGCAATCAAGCAAGAAAGT
DM29	466-505	ATG 488 mutagenesis	GGAGAAGGAGTATCGGTCCGGTCTGGATGACCTCACGTCA
DM31	475-498	ATG 488 mutagenesis	GTATCGGTCCGGTCATGGATGACCT
DM32	1196-1246	Acidic domain deletion	CTCCACCCTCGTCATC//TGTGGGGAATACTC
DM33	1382-1426	Glycine domain deletion	TCTTCCCACCTTCCCG//TACAGAAAGCAGCTTC
DM34	3216-3192	3204 mutagenesis	GCTCCAAAGGCTAGTTGAATCCTTC
DM35	3704-3727	3716 mutagenesis reverse	ACTTGGCGTCTCCGCAATGCTTAGA
DM35comp	3704-3727	3716 mutagenesis forward	TCTAAGCATTGCGGAGACGCCAAGT
DM36	4305-4329	4317 mutagenesis reverse	GCAGAAAGCTTCCCGCAGGTGCCCCGC
DM36comp	4305-4329	4317 mutagenesis forward	GCGGGCACCTGCGGGAAGCTTCTGC
DM37	3358-3381	3368 mutagenesis	ACATTACTACTCACGTGAGGGTCA
IB-bio	55-77	Frag I nested + biotin	CGCGGCCCTGGGTAGTCTGTAGA
IIB-bio	898-921	Frag II nested + biotin	GTATTTGGCAGATCAAGCAAAACT
IIIB-bio	1643-1665	Frag III nested + biotin	CACTGCCAGCAGGCAGGAGGAAAT
IVB-bio	2106-2129	Frag IV nested + biotin	CACCGATGCAAATAACCTCCGAG
VB-bio	2801-2824	Frag V nested + biotin	TTTGTGTTTCTTCTGACCACGCGG

VIB-bio	3814-3838	Frag VI nested + biotin	GAGCAATGACGATTATGTTTGGAA
VB2-bio	3098-3123	Frag Vb nested + biotin	TGACTAGTCCTGATGCATCCCAGAG
179no.1	31-54	Frag I - 5' PCR	TCTCTGTTTCCCTTGTTGGCGCTCG
179no.2	1142-1165	Frag I - 3' PCR	CTCTGGCCTCATGTCTGACTCCCA
176-25mer	868-892	Frag II - 5' PCR	CATGCTTAATGAAGCATCAGGCTTC
176-24-mer	1734-1757	Frag II - 3' PCR	TTGAACCAACGAGTCCTGATCTTGC
E6-5	1619-1641	Frag III - 5' PCR	GGTGTTAGGTGGCTGTGGGAATT
E6-6	2266-2290	Frag III - 3' PCR	GGTATCTCGTAAGACACATGCACAC
E6-1	2081-2104	Frag IV - 5' PCR	CCTCATCGGATCATTCTGTCTGGC
E6-2	3201-3227	Frag IV - 3' PCR	CATCATGGTCTGCTCCAAAGGCTGGTT
E6-3	2771-2797	Frag V - 5' PCR	ATTACGAGATACAAATGAGGACATCC
E6-4	4230-4253	Frag V - 3' PCR	TAGCCAAGAGTGAGGAGGAAGCGA
E6-9	3051-3075	Frag Vb - 5' PCR	AACAAAGGCGGTTTTCATAATCCAA
E6-10	3989-4011	Frag Vb - 3' PCR	ACACCAGACACTGCTCCCAGACA
179no.5	3786-3809	Frag VI - 5' PCR	GTGAAAACAAGAGTGAGGCCAAGG
179no.6	4689-4714	Frag VI - 3' PCR	CCTTCCTTAAAGTTTAAATTCAGAG

APPENDIX II

Alignment of the sequences of the human and mouse CSB proteins and the yeast homolog Rad26

1	M P N E G I P H S S Q T Q E Q D C L Q S	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
1	- - - - - - - - - - - - - H C - - -	SCRAD26.PRO
21	Q P V S N N E E M A I K Q E S G G D G E	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
3	- - - - - - - - - - - - - - - -	SCRAD26.PRO
41	V E E - L S F R S V G D G L S T S - V G	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
3	- - - - - - - - - - - - - - - -	SCRAD26.PRO
61	C A S A A F R R G F A L L H I D R H Q	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
3	- - - - S E K H T S V M E D K E Q Q D N	SCRAD26.PRO
81	Q A V E P S - Q - L E L G G L G V D V I	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
19	A K L E N N E - - - S K D L G V N V L	SCRAD26.PRO
101	D Q D V L E Q G V L Q G V D N A I H E -	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
36	S Q S S L E E K I A N D V T H F - - -	SCRAD26.PRO
121	S R A S G L V D V E - E - R S V L D D L	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
52	- - - S N L Q S L Q Q E - - - - E T R I	SCRAD26.PRO
141	T S C T T S L R Q - - - - E G L S P G	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
65	E R S K T A L Q R Y V N K K N H L T R K	SCRAD26.PRO
161	A A T S R D - N R - L D S V K R Q K - N	ERCC6.PRO
1	- - - - - - - - - - - - - - - -	MOCSB1.PRO
85	L N N T T R - S V K Q N L R D Q I K N L	SCRAD26.PRO
181	K E Q Q L K K - T - - Q K H L Q - - L G	ERCC6.PRO
1	- E Q Q L K K - T - - Q K R L C A - L G	MOCSB1.PRO
105	Q S D D I E R V L K D I D D I C S R I -	SCRAD26.PRO
201	G A E V - I E L D T - S L E E E - - E E P	ERCC6.PRO
20	G A G V Q V E L D H E S L F E E D D - E E P	MOCSB1.PRO
124	- K E L - E Q V L Q G A E N K G S K F G	SCRAD26.PRO
220	G P S S L G S M L M P V Q E T A W E E L	ERCC6.PRO
40	G P S C L G S M L M P A Q E T A W E F L	MOCSB1.PRO
143	- - - - - - - L Q R P G E T E K E F L	SCRAD26.PRO
240	I R T G Q N T F F G T G I F Q K Q E K K	ERCC6.PRO
60	I R T G Q N T F F G T P A P Q K Q E K K	MOCSB1.PRO
155	I R T G K I T A F G - - - - - - - -	SCRAD26.PRO
260	P R K I M L N E - S G F E K Y L A D Q A	ERCC6.PRO
80	P R K I M L N E A S G F E K Y L A E Q A	MOCSB1.PRO
165	- - - - - - - - - - - - - H K A	SCRAD26.PRO
280	K L S F E R K K Q G C H K R A A R K A P	ERCC6.PRO
100	Q L S F E R K K Q A A T K R T A K K A X	MOCSB1.PRO
168	G F S L D T A N R E Y A K - - - - -	SCRAD26.PRO
300	A P V T P P A P V Q N K N K P N K K A R	ERCC6.PRO
120	X X X X X X X X X X X X X X X X X	MOCSB1.PRO
181	- - - - - - - N D E Q K D E D F E M	SCRAD26.PRO
320	V L S K K E E R L K K H I K K L Q K R A	ERCC6.PRO
140	V L S Q T D N R L K K H S R K L Q R R A	MOCSB1.PRO
192	A T E Q M V E N L T D E D D N L S D Q D	SCRAD26.PRO

340	L	Q	F	Q	G	K	V	G	L	P	K	A	R	R	P	W	E	S	D	M	ERCC6.PRO
160	L	Q	F	Q	G	K	V	G	L	P	S	G	K	K	P	L	E	P	E	V	MOCSB1.PRO
212	Y	Q	M	S	G	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
360	R	P	E	A	E	G	D	S	E	G	E	E	S	E	Y	F	P	T	E	E	ERCC6.PRO
180	R	P	E	A	E	G	D	T	E	G	E	E	S	G	S	S	P	T	D	G	MOCSB1.PRO
218	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	S	SCRAD26.PRO
380	E	E	E	E	E	D	D	E	V	E	G	A	E	A	D	L	S	G	D	G	ERCC6.PRO
200	E	E	E	E	E	Q	E	E	E	E	G	V	-	A	S	L	S	S	D	D	MOCSB1.PRO
220	E	D	D	E	E	E	E	N	D	D	K	I	-	-	-	-	-	-	-	-	SCRAD26.PRO
400	T	D	V	E	L	K	P	L	P	K	G	G	K	R	Q	K	K	V	P	V	ERCC6.PRO
219	V	S	Y	E	L	K	P	L	R	K	Q	-	K	Y	Q	K	K	V	P	V	MOCSB1.PRO
232	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	L	SCRAD26.PRO
420	Q	E	I	D	D	D	F	F	F	S	S	G	E	E	A	E	A	A	S	V	ERCC6.PRO
238	Q	E	I	D	D	D	F	F	F	S	S	E	E	E	D	E	A	M	-	-	MOCSB1.PRO
233	K	E	L	E	D	L	R	F	R	G	Q	P	G	E	A	-	-	-	-	-	SCRAD26.PRO
440	G	E	G	G	G	G	G	R	K	V	G	R	-	R	D	D	G	D	E	D	ERCC6.PRO
256	-	E	G	R	G	G	G	R	K	V	A	R	R	Q	D	D	G	D	E	D	MOCSB1.PRO
248	-	-	-	-	-	-	-	-	-	-	-	-	-	K	D	D	G	D	E	L	SCRAD26.PRO
460	Y	Y	K	Q	R	L	R	R	W	N	K	L	R	L	Q	D	K	E	K	R	ERCC6.PRO
275	Y	Y	K	Q	R	L	R	X	X	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
255	Y	Y	Q	E	R	L	K	K	W	V	K	Q	R	S	C	G	S	Q	R	S	SCRAD26.PRO
480	L	K	L	E	D	-	-	-	-	D	S	E	E	S	D	A	F	F	D	E	ERCC6.PRO
295	X	X	X	X	X	-	-	-	-	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
275	S	D	L	P	E	W	R	R	P	H	P	N	I	P	D	A	K	L	N	S	SCRAD26.PRO
496	G	F	K	V	F	G	F	L	F	K	K	L	F	K	-	Q	Q	T	G	V	ERCC6.PRO
311	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
295	Q	F	K	I	P	G	E	I	Y	S	L	L	F	N	-	Q	K	T	C	V	SCRAD26.PRO
516	R	W	L	W	E	L	H	C	Q	Q	A	G	G	I	L	G	D	E	M	G	ERCC6.PRO
331	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
315	Q	W	L	Y	E	L	Y	Q	C	N	C	G	G	I	I	G	D	E	M	G	SCRAD26.PRO
536	L	G	K	T	I	Q	-	-	E	L	A	G	L	S	-	S	K	-	R	-	ERCC6.PRO
351	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
335	L	G	K	T	I	Q	V	-	L	F	I	A	A	L	H	H	S	G	L	L	SCRAD26.PRO
556	T	R	G	S	N	Y	R	F	E	G	L	G	F	T	V	I	V	C	P	T	ERCC6.PRO
371	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
355	T	-	-	-	-	-	-	-	-	-	-	G	F	V	L	I	V	C	P	A	SCRAD26.PRO
576	T	V	M	H	Q	W	V	K	E	-	F	H	T	W	W	P	P	F	R	V	ERCC6.PRO
391	X	X	X	X	X	X	X	X	E	F	L	S	W	W	P	P	F	R	V	-	MOCSB1.PRO
365	T	V	M	K	Q	W	C	N	E	-	F	H	H	W	W	P	P	L	R	T	SCRAD26.PRO
595	A	I	L	H	E	T	G	S	-	-	-	-	-	-	-	-	-	-	-	-	ERCC6.PRO
411	A	V	L	H	E	T	G	S	-	-	-	-	-	-	-	-	-	-	-	-	MOCSB1.PRO
384	V	I	L	H	S	M	G	S	G	M	A	S	D	Q	K	F	K	M	D	E	SCRAD26.PRO
603	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	T	H	K	ERCC6.PRO
419	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	T	H	K	MOCSB1.PRO
404	N	D	L	E	N	L	I	M	N	S	K	P	S	D	F	S	Y	E	D	W	SCRAD26.PRO
607	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	K	L	I	ERCC6.PRO
423	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	R	L	I	MOCSB1.PRO
424	K	N	S	T	R	T	K	K	A	L	E	S	S	Y	H	L	D	K	L	I	SCRAD26.PRO
612	R	D	V	A	H	C	H	G	I	L	I	T	S	Y	S	Y	I	R	L	M	ERCC6.PRO
428	R	D	I	V	Y	C	H	G	V	L	I	T	S	Y	S	Y	I	R	L	M	MOCSB1.PRO
444	D	K	V	V	T	D	G	H	I	L	I	T	T	Y	V	G	L	R	I	H	SCRAD26.PRO

632	Q D D I S R Y D W H Y V I L D E G H K I	ERCC6.PRO
448	Q D D I S R H D W H Y V I L D E G H K I	MOCSB1.PRO
464	S D K L L K V K W Q Y A V L D F G H K I	SCRAD26.PRO
652	R N P N A A V T L A C K Q F R T P H R I	ERCC6.PRO
468	R N P N A A V T L P C K Q F R T P H R I	MOCSB1.PRO
484	R N P D S E I S L T C K K L K T H N R	SCRAD26.PRO
672	I L S G S P M Q N N L R E L W S L F D F	ERCC6.PRO
488	I L S G S F M Q N N L R E L W S L F D F	MOCSB1.PRO
504	I L S G T P I Q N N L T E L W S L F D F	SCRAD26.PRO
692	I F P G K L G T L P V F M E Q F S V P	ERCC6.PRO
508	I F P G K L G T L P V F M E Q F S V P	MOCSB1.PRO
524	I F P G K L G T L P V F Q Q Q F V I P	SCRAD26.PRO
712	T M G G S N A S F V Q V T A I K C A	ERCC6.PRO
528	T M G G S N A S F V L V T A I K C A	MOCSB1.PRO
544	N I G G A N A T N I Q V Q T G I K C A	SCRAD26.PRO
732	C V L R D T N F L L R R M P S D V K	ERCC6.PRO
548	C V L R D T N F L L R R M P S D V K	MOCSB1.PRO
564	V A L R D L S F Y L L R R V A A D V A	SCRAD26.PRO
752	M S L S L P D K N E Q V L F C R L T D E	ERCC6.PRO
568	M S L S L P D K N E Q V L F C R L T D E	MOCSB1.PRO
584	K D L - - F Q K K E M V L F C K L T K Y	SCRAD26.PRO
772	Q H K V Q F V D S K E R I L N G	ERCC6.PRO
588	Q H K V Q F I D S K A V R I L N G	MOCSB1.PRO
602	Q R S K L E F L H S S D L N Q I Q N G	SCRAD26.PRO
792	E M Q I F S G L L R R M P S D V K	ERCC6.PRO
608	E N Q I F S G L V L R R M P S D V K	MOCSB1.PRO
622	K R N V L F G I D I L R R M P S D V K	SCRAD26.PRO
812	F S G G P K N L A S G P P E D E L E E D Q	ERCC6.PRO
628	F S G G P K N A S G P P E D E L E E E Q	MOCSB1.PRO
642	L D R D T K R - - - - - H N P D	SCRAD26.PRO
832	F G Y W K R S G M I V V E S L L W	ERCC6.PRO
648	F G H W R R S G M I V V E S L L W	MOCSB1.PRO
653	Y G D P K K S G M Q V V K Q L L L W	SCRAD26.PRO
852	H K Q G Q R V L L F S Q S R Q M L D L	ERCC6.PRO
668	H R Q G Q R V L L C S Q S R Q M L H L	MOCSB1.PRO
673	H K Q G Y K A L L F T Q S R Q M L D L	SCRAD26.PRO
872	E V F L R A Q K - - - - - T L A M D	ERCC6.PRO
688	E V F L R A H K - - - - - S L A M D	MOCSB1.PRO
693	E E F I S T K D P D L S H L N L R M D	SCRAD26.PRO
887	G T T T I A S R Q P L I T R N E D T S	ERCC6.PRO
703	G T T T I A S R Q P L I T R Y N E D T S	MOCSB1.PRO
713	G T T N I K G R Q S L V D R F N N E S	SCRAD26.PRO
907	I F V F L L T T R V G G L G V N L T G A	ERCC6.PRO
723	I F V F L L T T R V G G L G V N L T G A	MOCSB1.PRO
732	F D V F L L T T R V G G L G V N L T G A	SCRAD26.PRO
927	N R V V I D F D W N P S T D T Q A R E	ERCC6.PRO
743	N R V I I Y N P D W N P S T D T Q A R E	MOCSB1.PRO
752	N R I I I F D P D W N P S T D M Q A R E	SCRAD26.PRO
947	R A W R I G Q K K Q V T V R L L T A G	ERCC6.PRO
763	R A W R I G Q K K Q V T V R L L T A G	MOCSB1.PRO
772	R A W R I G Q K R E V S I Y R L M V G G	SCRAD26.PRO

967	T	I	E	E	K	I	Y	H	R	Q	I	F	K	Q	F	L	T	N	R	V	ERCC6.PRO
783	T	I	E	E	K	I	Y	H	R	Q	I	F	K	Q	F	L	T	N	R	V	MOCSB1.PRO
792	S	I	E	E	K	I	Y	H	R	Q	I	F	K	Q	F	L	T	N	R	I	SCRAD26.PRO
987	L	K	D	P	K	Q	R	R	F	F	K	S	N	D	L	Y	E	L	F	T	ERCC6.PRO
803	L	K	D	P	K	X	X	X	F	F	R	S	N	D	L	Y	E	L	F	T	MOCSB1.PRO
812	L	T	D	P	K	Q	K	R	F	F	K	I	H	E	L	H	D	L	F	S	SCRAD26.PRO
1007	L	T	S	P	D	A	S	Q	S	T	E	T	S	A	I	F	A	G	T	G	ERCC6.PRO
823	L	T	S	P	D	A	S	Q	G	T	E	T	S	A	I	R	A	G	T	G	MOCSB1.PRO
832	L	G	G	E	N	G	Y	S	T	E	E	L	N	-	-	-	-	-	-	-	SCRAD26.PRO
1027	S	D	V	Q	T	P	K	C	H	L	K	R	R	I	Q	P	A	F	G	A	ERCC6.PRO
843	S	D	I	Q	T	P	K	C	Q	L	K	K	R	T	S	T	V	L	G	T	MOCSB1.PRO
845	E	E	V	Q	K	H	T	E	N	L	K	N	S	-	-	-	-	-	-	-	SCRAD26.PRO
1047	D	H	D	V	P	K	R	K	K	F	P	A	S	N	I	S	V	N	D	A	ERCC6.PRO
863	D	-	-	-	P	K	C	K	K	P	P	V	S	D	T	P	A	N	A	A	MOCSB1.PRO
858	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
1067	T	S	S	E	E	K	S	E	A	K	G	A	E	V	N	A	V	T	S	N	ERCC6.PRO
880	T	L	I	G	E	K	P	K	A	A	G	A	T	G	R	S	V	T	S	G	MOCSB1.PRO
858	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	K	S	E	SCRAD26.PRO
1087	R	S	D	F	L	K	D	D	P	H	M	S	S	H	V	T	S	N	D	R	ERCC6.PRO
900	E	S	G	P	F	K	G	D	H	D	T	N	G	H	R	A	S	S	V	A	MOCSB1.PRO
861	E	S	D	D	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
1107	L	G	E	E	T	N	A	V	S	G	P	E	E	L	S	V	I	S	G	N	ERCC6.PRO
920	F	G	E	E	T	D	A	G	S	T	L	E	H	L	S	V	M	S	G	D	MOCSB1.PRO
866	-	-	-	-	-	-	-	-	-	-	-	E	Q	L	V	N	L	S	G	V	SCRAD26.PRO
1127	G	E	C	S	N	S	S	G	T	G	K	T	S	M	P	S	G	D	E	S	ERCC6.PRO
940	G	K	H	S	D	S	P	T	V	D	H	T	S	R	E	P	V	E	F	A	MOCSB1.PRO
875	S	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
1147	I	D	E	K	L	G	L	S	I	K	-	R	F	R	P	S	Q	-	Q	T	ERCC6.PRO
960	P	V	R	S	R	V	L	P	M	Q	G	R	A	A	D	C	Q	-	Q	T	MOCSB1.PRO
877	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
1166	E	A	F	W	E	N	F	Q	M	E	N	H	F	-	F	H	K	S	K	T	ERCC6.PRO
980	E	P	V	P	M	S	E	Q	M	E	G	Q	F	S	K	Y	K	S	K	R	MOCSB1.PRO
877	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
1186	K	H	H	S	V	A	E	E	E	T	L	E	F	H	L	R	P	K	Q	K	ERCC6.PRO
1000	K	H	D	-	-	A	S	E	X	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
877	-	-	-	-	-	-	-	-	-	-	L	F	S	F	Y	N	G	K	E	K	SCRAD26.PRO
1206	P	K	N	S	K	H	C	R	D	A	K	F	E	G	T	R	I	P	H	L	ERCC6.PRO
1018	X	K	N	S	K	H	C	R	D	A	K	F	E	G	T	R	V	P	H	L	MOCSB1.PRO
887	K	E	N	S	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
1226	V	K	K	R	R	Y	Q	K	Q	D	S	E	N	K	S	E	A	K	E	Q	ERCC6.PRO
1038	V	K	K	R	R	Y	R	Q	Q	T	S	E	Q	E	G	G	A	K	D	R	MOCSB1.PRO
892	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO
1246	S	N	D	D	Y	V	L	E	K	L	F	K	K	S	V	G	V	H	S	V	ERCC6.PRO
1058	S	S	D	D	Y	V	L	E	K	L	F	K	K	S	X	X	X	X	X	X	MOCSB1.PRO
892	T	E	D	D	R	L	I	E	G	L	L	G	G	E	S	N	L	E	T	V	SCRAD26.PRO
1266	M	K	H	D	A	I	M	D	G	-	-	A	S	P	D	Y	V	L	V	E	ERCC6.PRO
1078	X	X	X	X	X	X	X	X	X	-	-	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
912	M	S	H	D	S	V	V	N	S	H	A	G	S	S	S	S	N	I	I	T	SCRAD26.PRO
1284	A	E	A	N	R	V	A	Q	D	A	L	K	A	L	R	L	S	R	Q	R	ERCC6.PRO
1096	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	MOCSB1.PRO
932	K	E	A	S	R	V	A	I	E	A	V	N	A	L	R	K	S	R	K	K	SCRAD26.PRO

1304	C	L	G	A	V	S	-	G	V	F	T	W	T	G	H	R	G	I	S	G	ERCC6.PRO	
1116	X	L	G	A	A	S	-	G	V	F	T	W	T	G	H	R	G	I	S	G	MOCSB1.PRO	
952	I	T	K	Q	Y	E	I	G	T	P	T	W	T	G	SCRAD26.PRO	
1323	A	P	A	G	K	K	S	R	F	G	K	K	R	N	S	N	F	S	V	Q	ERCC6.PRO	
1135	A	P	T	G	V	K	N	R	F	G	Q	K	R	D	S	S	L	P	V	Q	MOCSB1.PRO	
966	-	-	-	-	-	-	-	R	F	G	-	-	-	-	-	-	-	-	-	-	SCRAD26.PRO	
1343	H	P	S	S	T	S	P	T	E	K	C	Q	D	G	I	M	K	K	E	G	ERCC6.PRO	
1155	H	P	S	S	L	-	-	T	E	K	T	Q	N	N	-	M	K	K	F	G	MOCSB1.PRO	
969	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	K	A	G	SCRAD26.PRO	
1363	K	D	N	V	P	E	H	F	S	G	R	A	E	D	A	D	S	S	S	G	ERCC6.PRO	
1172	K	A	H	T	P	E	H	F	S	G	K	-	E	D	G	A	S	V	S	G	MOCSB1.PRO	
972	E	I	R	K	R	D	P	L	K	N	K	L	T	G	S	A	A	I	L	G	SCRAD26.PRO	
1383	P	L	A	S	S	S	L	L	A	K	M	R	A	R	N	H	L	-	L	P	ERCC6.PRO	
1191	A	P	S	S	S	S	L	L	A	R	M	R	A	R	N	H	M	-	L	P	MOCSB1.PRO	
992	N	I	T	K	S	Q	K	E	-	S	K	E	A	R	Q	E	N	Y	-	-	SCRAD26.PRO	
1403	E	R	L	E	S	E	S	G	H	L	Q	E	A	S	A	L	L	F	-	T	ERCC6.PRO	
1211	E	R	L	E	S	D	S	E	H	L	A	E	A	A	V	P	F	C	G	-	MOCSB1.PRO	
1010	-	-	-	-	D	D	G	I	T	F	A	R	S	K	E	I	-	-	-	-	SCRAD26.PRO	
1422	T	E	H	D	D	L	L	V	E	N	R	N	F	I	A	F	Q	A	H	T	ERCC6.PRO	
1231	T	E	H	D	D	L	L	V	D	M	R	N	F	I	A	F	Q	A	Q	V	MOCSB1.PRO	
1022	N	S	N	T	K	T	L	E	N	I	R	A	Y	L	-	-	Q	K	Q	N	SCRAD26.PRO	
1442	D	G	Q	A	S	T	R	E	-	L	Q	E	F	E	S	K	L	S	A	S	ERCC6.PRO	
1251	D	G	Q	A	S	T	Q	E	-	L	Q	-	-	-	-	-	-	-	-	-	MOCSB1.PRO	
1040	N	F	F	S	S	S	V	S	-	L	N	S	I	G	V	S	L	S	D	K	SCRAD26.PRO	
1462	Q	S	C	V	-	-	F	R	E	L	L	R	N	L	C	T	F	H	R	T	S	ERCC6.PRO
1261	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	MOCSB1.PRO	
1060	E	D	V	I	K	V	R	A	L	L	K	T	I	A	Q	F	D	K	E	R	SCRAD26.PRO	
1481	G	G	E	G	I	W	K	L	K	P	E	Y	-	-	-	-	-	-	-	C	ERCC6.PRO	
1261	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	MOCSB1.PRO	
1080	K	G	-	-	-	-	W	V	L	D	E	E	F	R	N	N	N	A	S	-	SCRAD26.PRO	

Black boxes indicate residues identical to the human CSB sequence. The mouse sequence is incomplete and X indicates sequence that is unavailable.

ERCC6 (CSB) sequence (Troelstra et al. 1992). Mouse CSB (van der Horst, Mallery unpublished data, EST accession no. AA675285). Rad26 (van Gool et al. 1994).

APPENDIX III

The full nucleotide sequence and amino acid translation of *CSB*.

TGGGTTCCAAGGCGGCTGGCGGCGGTAGCGTCTCTGTTTCCTTGTGGGCGCTCGCGCGGC	60
<hr/>	
CCTGGGTAGTCTGTAGAGAATGCCAAATGAGGGAATCCCCCACTCAAGTCAAACCTCAGGA	120
<hr/>	
M P N E G I P H S S Q T Q E	
GCAAGACTGTTTACAGAGTCAACCTGTCAGTAATAATGAAGAAATGGCAATCAAGCAAGA	180
<hr/>	
Q D C L Q S Q P V S N N E E M A I K Q E	
AAGTGGTGGTGATGGGGAGGTGGAGGAGTACCTGTCCTTTCGTTCTGTGGGTGACGGGCT	240
<hr/>	
S G G D G E V E E Y L S F R S V G D G L	
GTCCACCTCTGCTGTGGGGTGCGCATCAGCAGCTCCGAGGAGAGGGCCAGCCCTGCTGCA	300
<hr/>	
S T S A V G C A S A A P R R G P A L L H	
CATCGACCGACATCAGATCCAGGCAGTAGAGCCTAGCGCCCAGGCCCTTGAGCTGCAGGG	360
<hr/>	
I D R H Q I Q A V E P S A Q A L E L Q G	
TTTGGGTGTGGACGTCTATGACCAGGACGTGCTGGAACAGGGAGTGCTTCAGCAGGTGGA	420
<hr/>	
L G V D V Y D Q D V L E Q G V L Q Q V D	
CAATGCCATCCATGAGGCCAGCCGTGCCTCCCAGCTCGTTGACGTGGAGAAGGAGTATCG	480
<hr/>	
N A I H E A S R A S Q L V D V E K E Y R	
GTCGGTCCTGGATGACCTCACGTGATGTACGACATCCCTAAGGCAAATCAATAAAATTAT	540
<hr/>	
S V L D D L T S C T T S L R Q I N K I I	
TGAACAGCTTAGCCCTCAAGCTGCCACCAGCAGAGACATCAACAGGAACTAGATTCTGT	600
<hr/>	
E Q L S P Q A A T S R D I N R K L D S V	
AAAACGACAGAAGTATAATAAGGAACAACAGCTAAAAAAGATCACTGCAAAACAAAAGCA	660
<hr/>	
K R Q K Y N K E Q Q L K K I T A K Q K H	
TCTCCAGGCCATCCTTGGAGGAGCAGAGGTGAAAATTGAACTAGATCACGCCAGTCTGGA	720
<hr/>	
L Q A I L G G A E V K I E L D H A S L E	
GGAGGATGCAGAGCCGGGGCCATCCAGTCTTGGCAGCATGCTCATGCCTGTCCAGGAGAC	780
<hr/>	
E D A E P G P S S L G S M L M P V Q E T	

TGCCTGGGAAGAGCTCATCCGCACTGGCCAGATGACACCTTTTGGTACCCAGATCCCTCA 840
A W E E L I R T G Q M T P F G T Q I P Q
GAAACAGGAGAAAAAGCCCAGAAAAATCATGCTTAATGAAGCATCAGGCTTCGAAAAGTA 900
K Q E K K P R K I M L N E A S G F E K Y
TTTGGCAGATCAAGCAAAACTGTCTTTTGAAAGGAAGAAGCAAGGTTGTAATAAAAGAGC 960
L A D Q A K L S F E R K K Q G C N K R A
AGCTAGAAAAGCTCCAGCCCCAGTCACGCCTCCAGCCCCAGTGCAAAATAAAAACAAACC 1020
A R K A P A P V T P P A P V Q N K N K P
AAACAAGAAAGCCAGAGTTCTGTCCAAAAAAGAGGAGCGTTTGAAAAAGCACATCAAGAA 1080
N K K A R V L S K K E E R L K K H I K K
ACTCCAGAAGAGGGCTTTGCAGTTCCAGGGGAAAGTGGGATTGCCAAAGGCAAGGAGACC 1140
L Q K R A L Q F Q G K V G L P K A R R P
TTGGGAGTCAGACATGAGGCCAGAGGCAGAGGGAGACTCTGAGGGTGAAGAGTCTGAGTA 1200
W E S D M R P E A E G D S E G E E S E Y
TTTCCCCACAGAGGAGGAGGAAGAGGAGGAAGATGACGAGGTGGAGGGGGCAGAGGCGGA 1260
F P T E E E E E E E D D E V E G A E A D
CCTGTCTGGAGATGGTACTGACTATGAGCTGAAGCCTCTGCCCAAGGGCGGGAAACGGCA 1320
L S G D G T D Y E L K P L P K G G K R Q
GAAGAAAGTGCCAGTGCAGGAGATTGATGATGACTTTTTTCCAAGTTCTGGGGAAGAAGC 1380
K K V P V Q E I D D D F F P S S G E E A
TGAAGCTGCTTCTGTAGGAGAAGGAGGAGGAGGAGGTCGGAAAGTGGGAAGATACCGAGA 1440
E A A S V G E G G G G G R K V G R Y R D
TGATGGAGATGAAGATTATTATAAGCAGCGGTTAAGGAGATGGAATAAACTGAGACTGCA 1500
D G D E D Y Y K Q R L R R W N K L R L Q
GGACAAAGAGAAACGTCTGAAGCTGGAGGACGATTCTGAGGAAAGTGATGCTGAATTTGA 1560
D K E K R L K L E D D S E E S D A E F D

CGAAGGTTTTAAAGTGCCAGGTTTTCTGTTCAAAAAGCTTTTTAAGTACCAGCAGACAGG 1620
E G F K V P G F L F K K L F K Y Q Q T G
TGTTAGGTGGCTGTGGGAATTGCACTGCCAGCAGGCAGGAGGAATTCTGGGAGATGAAAT 1680
V R W L W E L H C Q Q A G G I L G D E M
GGGATTGGGCAAGACCATCCAGATAATTGCCTTCTTGGCAGGTCTGAGCTACAGCAAGAT 1740
G L G K T I Q I I A F L A G L S Y S K I
CAGGACTCGTGGTTCAAATTACAGGTTTGAGGGGTTGGGTCCAACCTGTAATTGTCTGTCC 1800
R T R G S N Y R F E G L G P T V I V C P
AACAACAGTGATGCATCAGTGGGTGAAGGAATTTACACGTGGTGGCCTCCGTTTCAGAGT 1860
T T V M H Q W V K E F H T W W P P F R V
GGCAATTCTACATGAAACCGGTTCTATACCCACAAAAAGGAGAACTAATTCGAGATGT 1920
A I L H E T G S Y T H K K E K L I R D V
TGCTCATTGTCATGGAATTTTGATCACATCTTACTCCTACATTGATTGATGCAGGATGA 1980
A H C H G I L I T S Y S Y I R L M Q D D
CATTAGCAGGTATGACTGGCACTATGTGATCTTGGACGAAGGACACAAAATTCGAAATCC 2040
I S R Y D W H Y V I L D E G H K I R N P
AAATGCTGCTGTCACCCTTGCTTGCAAACAGTTTCGCACCCCTCATCGGATCATTCTGTC 2100
N A A V T L A C K Q F R T P H R I I L S
TGGCTCACCGATGCAAAATAACCTCCGAGAGCTGTGGTCGCTCTTTGACTTCATCTTCCC 2160
G S P M Q N N L R E L W S L F D F I F P
GGGAAAGTTAGGCACGTTGCCTGTGTTTATGGAGCAGTTCTCCGTCCCCATCACCATGGG 2220
G K L G T L P V F M E Q F S V P I T M G
GGGATATTCAAATGCTTCCCCAGTACAGGTCAAACTGCTTACAAGTGTGCATGTGTCTT 2280
G Y S N A S P V Q V K T A Y K C A C V L
ACGAGATACCATAAATCCATACCTACTGCGGAGAATGAAGTCAGATGTCAAGATGAGCCT 2340
R D T I N P Y L L R R M K S D V K M S L

TTCTTTGCCAGATAAAAATGAACAGGTCTTATTTTGCCGTCTTACAGATGAGCAGCATAA 2400
S L P D K N E Q V L F C R L T D E Q H K
AGTCTACCAAATTTTCGTTGATTCCAAAGAAGTTTACAGGATTCTCAATGGAGAGATGCA 2460
V Y Q N F V D S K E V Y R I L N G E M Q
GATTTTCTCCGGACTTATAGCCCTAAGAAAAATTTGCAACCACCCTGATCTCTTTTCTGG 2520
I F S G L I A L R K I C N H P D L F S G
AGGTCCCAAGAATCTCAAAGGTCTTCCTGATGATGAACTAGAAGAAGATCAGTTTGGGTA 2580
G P K N L K G L P D D E L E E D Q F G Y
CTGGAAACGTTCTGGGAAAATGATTGTTGTTGAGTCTTTGTTGAAAATATGGCACAAGCA 2640
W K R S G K M I V V E S L L K I W H K Q
GGGTCAGCGAGTATTGCTGTTTTCTCAGTCAAGGCAGATGCTGGACATACTTGAAGTATT 2700
G Q R V L L F S Q S R Q M L D I L E V F
CCTTAGAGCCCAAAGTATACCTATCTCAAGATGGATGGTACCACTACAATAGCTTCAAG 2760
L R A Q K Y T Y L K M D G T T T I A S R
ACAGCCACTGATTACGAGATACAATGAGGACACATCCATATTTGTGTTTCTTCTGACCAC 2820
Q P L I T R Y N E D T S I F V F L L T T
GCGGGTGGGCGGCTTAGGTGTCAACCTGACGGGGGCAAACAGAGTTGTCATCTATGACCC 2880
R V G G L G V N L T G A N R V V I Y D P
AGACTGGAACCCAAGCACGGACACGCAGGCCCGGGAGCGAGCATGGAGAATAGGCCAGAA 2940
D W N P S T D T Q A R E R A W R I G Q K
GAAGCAAGTGACTGTGTACAGGCTCCTGACTGCGGGCACCATTGAAGAAAAGATCTACCA 3000
K Q V T V Y R L L T A G T I E E K I Y H
CCGACAAATCTTCAAGCAGTTTTTTGACAAATAGAGTGCTAAAAGACCCAAAACAAAGGCG 3060
R Q I F K Q F L T N R V L K D P K Q R R
GTTTTTCAAATCCAATGATCTCTATGAGCTATTTACTCTGACTAGTCCTGATGCATCCCA 3120
F F K S N D L Y E L F T L T S P D A S Q

GAGCACTGAAACAAGTGCAATTTTTGCAGGAAGTGGATCAGATGTTTCAGACACCCAAATG 3180
S T E T S A I F A G T G S D V Q T P K C
CCATCTAAAAAGAAGGATTCAACCAGCCTTTGGAGCAGACCATGATGTTCCAAAACGCAA 3240
H L K R R I Q P A F G A D H D V P K R K
GAAGTTCCTGCTTCTAACATATCTGTAAATGATGCCACATCATCTGAAGAGAAATCTGA 3300
K F P A S N I S V N D A T S S E E K S E
GGCTAAAGGAGCTGAAGTAAATGCAGTAACTTCTAATCGAAGTGATCCTTTGAAAGATGA 3360
A K G A E V N A V T S N R S D P L K D D
CCCTCACATGAGTAGTAATGTAAGTAGCAATGATAGGCTTGGAGAAGAGACAAATGCAGT 3420
P H M S S N V T S N D R L G E E T N A V
ATCTGGACCAGAAGAGTTGTCAGTGATTAGTGGAATGGGGAATGTTCAAATTCTTCAGG 3480
S G P E E L S V I S G N G E C S N S S G
AACAGGCAAACTTCTATGCCATCTGGTGATGAAAGCATTGATGAAAAGTTAGGTCTTTC 3540
T G K T S M P S G D E S I D E K L G L S
TTACAAAAGAGAAAGACCCAGCCAGGCTCAAACAGAAGCTTTTTGGGAGAATAAACAAT 3600
Y K R E R P S Q A Q T E A F W E N K Q M
GGAAAATAATTTTTATAAGCACAAGTCAAAAACAAAACATCATAGTGTGGCAGAAGAAGA 3660
E N N F Y K H K S K T K H H S V A E E E
GACCCTGGAGAAACATCTGAGACCAAAGCAAAAGCCTAAGAACTCTAAGCATTGCAGAGA 3720
T L E K H L R P K Q K P K N S K H C R D
CGCCAAGTTTGAAGGAACTCGAATTCCACACCTGGTGAAGAAAAGGCGTTACCAGAAGCA 3780
A K F E G T R I P H L V K K R R Y Q K Q
AGACAGTGAAAACAAGAGTGAGGCCAAGGAACAGAGCAATGACGATTATGTTTTGGAAAA 3840
D S E N K S E A K E Q S N D D Y V L E K
GCTTTTCAAAAAATCAGTTGGCGTGCACAGTGTCATGAAGCACGATGCCATCATGGATGG 3900
L F K K S V G V H S V M K H D A I M D G

AGCCAGCCCAGATTATGTACTGGTGGAGGCAGAAGCCAACCGAGTGGCCCAGGATGCCCT	3960
A S P D Y V L V E A E A N R V A Q D A L	
GAAAGCACTGAGGCTCTCTCGTCAGCGGTGTCTGGGAGCAGTGTCTGGTGTTCACCTG	4020
K A L R L S R Q R C L G A V S G V P T W	
GACTGGCCACAGGGGGATTCTGGTGCACCAGCAGGAAAAAGAGTAGATTTGGTAAGAA	4080
T G H R G I S G A P A G K K S R F G K K	
AAGGAATTCTAACTTCTCTGTGCAGCATCCTTCATCAACATCTCCAACAGAGAAGTGCCA	4140
R N S N F S V Q H P S S T S P T E K C Q	
GGATGGCATCATGAAAAAGGAGGGGAAAAGATAATGTCCCTGAGCATTTTAGTGGAAGAGC	4200
D G I M K K E G K D N V P E H F S G R A	
AGAAGATGCAGACTCTTCATCCGGGGCCCCTCGCTTCCTCCTCACTCTTGGCTAAAATGAG	4260
E D A D S S S G P L A S S S L L A K M R	
AGCTAGAAACCACCTGATTCTGCCAGAGCGTTTAGAAAGTGAAAGCGGGCACCTGCAGGA	4320
A R N H L I L P E R L E S E S G H L Q E	
AGCTTCTGCCCTGCTGCCCACCACAGAACACGATGACCTTCTGGTGGAGATGAGAACTT	4380
A S A L L P T T E H D D L L V E M R N F	
CATCGCTTTCCAGGCCACACTGATGGCCAGGCCAGCACCAGGGAGATACTGCAGGAGTT	4440
I A F Q A H T D G Q A S T R E I L Q E F	
TGAATCCAAGTTATCTGCATCACAGTCTTGTGTCTTCCGAGAACTATTGAGAAATCTGTG	4500
E S K L S A S Q S C V F R E L L R N L C	
CACTTTCCATAGAACTTCTGGTGGTGAAGGAATTTGGAACTCAAGCCAGAATACTGCTA	4560
T F H R T S G G E G I W K L K P E Y C	
AACAACATTGCTTCCTAAACTTTCAAGTCCCTTTTCTAACGGGCATTTCTGATTATTAA	4620
TTTATTATTAATAATCATGTTTGTCAATGGAAGTTGGCTGCACTTGATGTTTGTTCAT	4680

GATGTCTACCTCAGAATTAAACTTTAAGGAAGG
→ 4714

The location and nature of the putative functional domains in *CSB* (all numbers are inclusive).

Functional Domain	Nucleotide position	Amino acid position
Acidic domain	1169-1261	363-394
Glycine stretch	1397-1417	439-446
Nuclear localisation signals	1475-1522 3191-3244	465-481 1037-1055
Casein kinase II phosphorylation site	1526-1549 3281-3301	482-490 1067-1074
Nucleotide binding site	3479-3493	1133-1138
Helicase domain I	1658-1711	526-544
Helicase domain Ia	1784-1822	568-581
Helicase domain II	1994-2026	638-649
Helicase domain III	2087-2143	669-688
Helicase domain IV	2651-2686	857-869
Helicase domain V	2789-2869	903-930
Helicase domain VI	2888-2929	936-950

The location and size of the twenty introns within the *CSB* genomic DNA.

Intron	Location	Size (kilobases)
1	65/66	6
2	501/502	1.8
3	622/623	2.1
4	731/732	4
5	1476/1477	19
6	1605/1606	5
7	1764/1765	7.5
8	1900/1901	9
9	2071/2072	0.5
10	2248/2249	3.5
11	2365/2366	2
12	2461/2462	2
13	2677/2678	0.5
14	2788/2789	0.4
15	2908/2909	0.4
16	3003/3004	1.1
17	3149/3150	0.085
18	3857/3858	8.5
19	4062/4063	0.8
20	4141/4142	1.2

Molecular Analysis of Mutations in the CSB (ERCC6) Gene in Patients with Cockayne Syndrome

Donna L. Mallery,¹ Bianca Tanganelli,² Stefano Colella,² Herdis Steingrimsdottir,^{1,*} Alain J. van Gool,^{3,†} Christine Troelstra,³ Miria Stefanini,² and Alan R. Lehmann¹

¹MRC Cell Mutation Unit, Sussex University, Falmer, Brighton; ²Istituto di Genetica Biochimica ed Evoluzionistica CNR, Pavia, Italy; and ³Department of Cell Biology and Genetics, Erasmus University, Rotterdam

Summary

Cockayne syndrome is a multisystem sun-sensitive genetic disorder associated with a specific defect in the ability to perform transcription-coupled repair of active genes after UV irradiation. Two complementation groups (CS-A and CS-B) have been identified, and 80% of patients have been assigned to the CS-B complementation group. We have analyzed the sites of the mutations in the CSB gene in 16 patients, to determine the spectrum of mutations in this gene and to see whether the nature of the mutation correlates with the type and severity of the clinical symptoms. In nine of the patients, the mutations resulted in truncated products in both alleles, whereas, in the other seven, at least one allele contained a single amino acid change. The latter mutations were confined to the C-terminal two-thirds of the protein and were shown to be inactivating by their failure to restore UV-irradiation resistance to hamster UV61 cells, which are known to be defective in the CSB gene. Neither the site nor the nature of the mutation correlated with the severity of the clinical features. Severe truncations were found in different patients with either classical or early-onset forms of the disease.

Introduction

The process of nucleotide excision repair (NER) protects cells from damage produced in cellular DNA by a wide

variety of carcinogens, including UV light. Three genetic disorders—xeroderma pigmentosum (XP), Cockayne syndrome (CS), and trichothiodystrophy (TTD)—are associated with defects in NER (Lehmann 1995). Most patients with XP are deficient in NER and contain mutations in one of seven genes (XPA–XPG) whose products are directly involved in NER. The recent discovery that the products of the XPB and XPD genes are components of the basal transcription factor TFIID—which has a dual role, in transcription and as a component of NER—revealed an unexpected link between DNA repair and transcription (reviewed in Lehmann 1995; Hoeijmakers et al. 1996).

The clinical features of XP—namely, multiple skin abnormalities, including a greatly elevated skin-cancer incidence caused by exposure to sunlight—are consistent with predictions based on the hypersensitivity and hypermutability of cultured XP cells to UV light. In contrast, CS patients do not have an increased frequency of skin cancers. CS is a multisystem disorder characterized by severe physical and mental retardation, microcephaly, progressive neurological and retinal degeneration, skeletal abnormalities, gait defects, and sun sensitivity but no increased frequency of cancer (Nance and Berry 1992). Like XP, CS cells are hypersensitive to the lethal effects of UV light, but NER of the bulk of genomic DNA is unaffected. CS cells are, however, defective in a subpathway of NER, known as “transcription-coupled repair” (TCR), whereby damage in the transcribed strand of active genes is rapidly and preferentially repaired. In CS cells, damage in active genes is repaired at the same relatively slow rate as the bulk DNA (van Hoffen et al. 1993). This correlates with earlier findings that RNA synthesis recovers rapidly after UV irradiation of normal cells, whereas this recovery does not occur in CS cells (Mayne and Lehmann 1982). Cell-fusion studies have assigned some 30 CS patients to two complementation groups, CS-A and CS-B (Tanaka et al. 1981; Lehmann 1982; Stefanini et al. 1996); 80% of them are in the CS-B group. In addition, there are a few rare cases of individuals with the clinical and cellular features of

Received August 28, 1997; accepted for publication November 5, 1997; electronically published January 16, 1998.

Address for correspondence and reprints: Dr. Alan R. Lehmann, MRC Cell Mutation Unit, Sussex University, Falmer, Brighton BN1 9RR, United Kingdom. E-mail: a.r.lehmann@sussex.ac.uk

*Present affiliation: Genpak, Falmer, Brighton.

†Present address: ICRF Clare Hall Laboratories, South Mimms, Porters Bar, United Kingdom.

© 1998 by The American Society of Human Genetics. All rights reserved.
0002-9297/98/6201-0014\$02.00

Table 1

Primers Used for Reverse Transcription and PCR

Primer	Sequence	Position ^a	Annealing Temperature
1F	5'-TCTCTGTTTCCTTGTGGGCGCTCG-3'	31-54	
1R	5'-CTCTGGCCTCATGTCTGACTCCCA-3'	1142-1165	65°C
2F	5'-CATGCTTAATGAAGCATCAGGCTTC-3'	868-892	
2R	5'-TTGAACCACGAGTCCTGATCTTGC-3'	1734-1757	63°C
3F	5'-GGTGTTAGGTGGCTGTGGGAATT-3'	1619-1641	
3R	5'-GGTATCTCGTAAGACACATGCACAC-3'	2266-2290	67°C
4F	5'-CCTCATCGGATCATTCTGTCTGGC-3'	2081-2104	
4R	5'-CATCATGGTCTGCTCCAAAGGCTGGTT-3'	3201-3227	67°C
5F	5'-ATTACGAGATACAATGAGGACACATCC-3'	2771-2805	
5R	5'-TAGCCAAGAGTGAGGAGGAAGCGA-3'	4237-4253	67°C
6F	5'-GTGAAAACAAGAGTGAGGCCAAGG-3'	3786-3809	
6R	5'-CCTTCCTTAAAGTTTAAATTCTGAGG-3'	4688-4714	60°C

^a The ORF starts at nucleotide 79 and ends at nucleotide 4558 (Troelstra et al. 1992).

both CS and XP. The defects in these individuals have been assigned to the XP-B, XP-D, and XP-G groups (Hoeijmakers 1993).

The genes defective in the CS-A (Henning et al. 1995) and CS-B (Troelstra et al. 1992) complementation groups have both been cloned recently. A gene designated "ERCC6" was cloned by its ability to correct the UV sensitivity of a mutant hamster cell line, UV61 (Troelstra et al. 1990). When the ERCC6 cDNA was introduced into different human cell lines, it was able to correct the UV sensitivity and post-UV-irradiation RNA-synthesis defect in a CS-B cell line, CS1AN (Troelstra et al. 1992), and this cell strain was found to contain mutations in the ERCC6 gene. ERCC6 was therefore subsequently renamed "CSB" (Lehmann et al. 1994). CSB encodes a 1,493-amino-acid protein, which contains seven domains characteristic of DNA helicases (Troelstra et al. 1992). However, the CSB protein is a member of the SWI2/SNF2 family of ATPases, and, like other members of this family, CSB does not appear to be a helicase (Selby and Sancar 1997). The function of the SWI/SNF protein family is thought to involve remodeling of protein-DNA interactions—such as chromatin structure—in different circumstances (Peterson and Tamkun 1995; Cairns et al. 1996; Pazin and Kadonaga 1997). The CSA gene has been cloned directly by its ability to correct the UV sensitivity of a CS-A cell line, CS3BE, and encodes a "WD-repeat" protein (Henning et al. 1995). In order to gain more insight into genotype-phenotype relationships in CS, we have identified mutations in the CSB gene in 16 CS patients whom we previously had assigned to this complementation group.

Material and Methods

Cell Culture

All CS cells were primary fibroblast cultures derived from patients with CS. Cultures were grown in Eagle's

minimal essential medium supplemented with 15% FCS. The UV-sensitive Chinese hamster cell line UV61 was grown in Eagle's minimal essential medium with 10% FCS.

cDNA Synthesis and PCR

RNA was extracted from $\sim 10^7$ cells, by use of lysis either (a) in NP40, by spinning out the nuclei and phenol extraction of the cytoplasmic fraction, followed by ethanol precipitation, or (b) with guanidinium isothiocyanate, followed by phenol extraction and isopropanol precipitation. cDNA synthesis was performed in three parts (A-C), by use of primers 2R, 4R, and 6R (see table 1). Reactions were performed by use of 1–5 μ g RNA and Moloney murine-leukemia-virus reverse transcriptase in a total volume of 20 μ l. After incubation for 1 h, the mixture was diluted to 40 μ l, and 5 μ l was used in PCR. Each of the cDNA samples was amplified in two parts, by the primers shown in table 1, so that the whole open reading frame (ORF) was amplified as six fragments. PCR mixtures contained 5 μ l appropriate cDNA, 20 pmol primers, and 0.125 mM each dNTP, in a total volume of 50 μ l. PCR conditions were 35 cycles at 94°C for 1 min, annealing (at temperatures indicated in table 1) for 1.5 min, and elongation at 72°C for 3 min. The products were reamplified by means of the same 3' primer and a nested biotinylated 5' primer, under conditions identical to those described above. PCR products were captured on streptavidin-coated Dynabeads and were sequenced by means of the dideoxy termination procedure with T7 DNA polymerase and ³²S-dATP, as described elsewhere (Steingrimsdottir et al. 1993).

In compound-heterozygote patients, in whom the mutations in the two alleles were different, the sequencing gel had both normal and mutant bands at the site of the mutation. In many of these cases, we confirmed the mutation by cloning the PCR products into a "T vector" and sequencing several clones.

Table 2

Mutations in CS-B Patients

Cell Line	Inactivating Mutation(s) ^a	Amino Acid Change	Location
CS2TAN	G1630A	Trp517stop	Exon 7
CS1TAN	C2282T	Arg735stop	Exon 11
CS8BR	C2639T	Gln854stop	Exon 13
25627	C2282T	Arg735stop	Exon 11
	C1436T	Arg453stop	Exon 5
CS4BR	C629T	Gln184stop	Exon 4
	C2087T	Arg670Trp	Exon 10
CS1BO	-G 1597 and C2918T	Frameshift 506→542stop and Arg947stop	Exon 6 and 16
	G3363C	Pro1095Arg	Exon 18
CS1BE	C2087T	Arg670Trp	Exon 10
	-A 3615	Frameshift 1179→1200stop	Exon 18
CS10BR	+26nt 3686	Frameshift 1203→1235stop	Exon 18
CS1ABR	-G 2223	Frameshift 715→738stop	Exon 10
CS10LO	+T 1359	Frameshift 427→435stop	Exon 5
CS2BE	-T 3614	Frameshift 1179→1200stop	Exon 18
	Deletion of exon 10	Deletion of 665-723	Exon 10
CS2BI	C2087T	Arg670Trp	Exon 10
	C3204T	Pro1042Leu	Exon 18
CS3TAN	T2630A	Trp851Arg	Exon 13
CS7TAN	A3716G	Arg1213Gly	Exon 18
CS1IAF	T2949G	Val957Gly	Exon 16
CS1MA	Deletion of exon 10	Del 665-723	Exon 10

^a Missense mutations shown to be inactivating by failure to restore UV resistance to UV61 cells.

Functional Complementation

Mutations were introduced into the CSB gene by means of two different procedures. PCR products containing the mutation of interest were cloned into a T vector and were sequenced to check for the absence of any PCR-induced spurious mutations. An appropriate restriction fragment containing the mutation was excised and used to replace the corresponding wild-type fragment in the intact CSB gene in BlueScript. Alternatively, the mutation was introduced directly into the gene by site-directed mutagenesis using the procedure of Kunkel et al. (1987). The entire ORF was checked for the absence of spurious mutations. The CSB gene was then transferred as a *Bam*HI fragment into the mammalian expression vectors pCDNA3 or pCIneo (Invitrogen). For transfection experiments, 5×10^5 UV61 cells were plated on 9-cm dishes. Two days later, 40 μ l polybrene was added to the plates, followed by 5 μ g DNA. After 6 h, the cells were treated with 20% dimethyl sulfoxide in serum-free medium for 4 min. Two days later, the cells were trypsinized, split into three aliquots, and replated in the presence of 1 mg G418/ml. The cells on one dish were grown and frozen as a mass population, and the second dish was subjected to UV-irradiation selection with three doses of 4 Jm⁻² with 2-day intervals between; individual colonies were isolated from the third plate and were grown for UV-irradiation-survival experiments.

Results

The 16 patients whom we have analyzed all failed to restore RNA synthesis to normal levels after UV irradiation, a phenotype that we regard as diagnostic for CS (Lehmann et al. 1993). Cell-fusion studies have assigned these patients to the CS-B group (Lehmann 1982; Stefanini et al. 1996). For molecular analysis, total cellular RNA was reverse transcribed by use of three CSB-specific primers. Each of the three cDNA samples was then amplified in two parts by PCR, so that the whole CSB cDNA was amplified in a total of six overlapping fragments of ~1 kb each, which were then sequenced directly. Results of our analyses are summarized in table 2 and figure 1.

Nonsense Mutations

Four cell strains contained nonsense mutations in both alleles, which resulted in truncation of >40% of the protein. Two Turkish patients, CS2TAN and CS1TAN, and one British patient of Asian origin, CS8BR, all the offspring of consanguineous marriages, were homozygous for G1630A, C2282T, and C2639T, respectively, which convert Trp517, Arg735, and Gln854, respectively, to stop codons. In 1978 we reported on a sun-sensitive patient, designated "11961" (Arlett et al. 1978), who was subsequently diagnosed as having CS and who was assigned to group B (Lehmann 1982). We have deter-

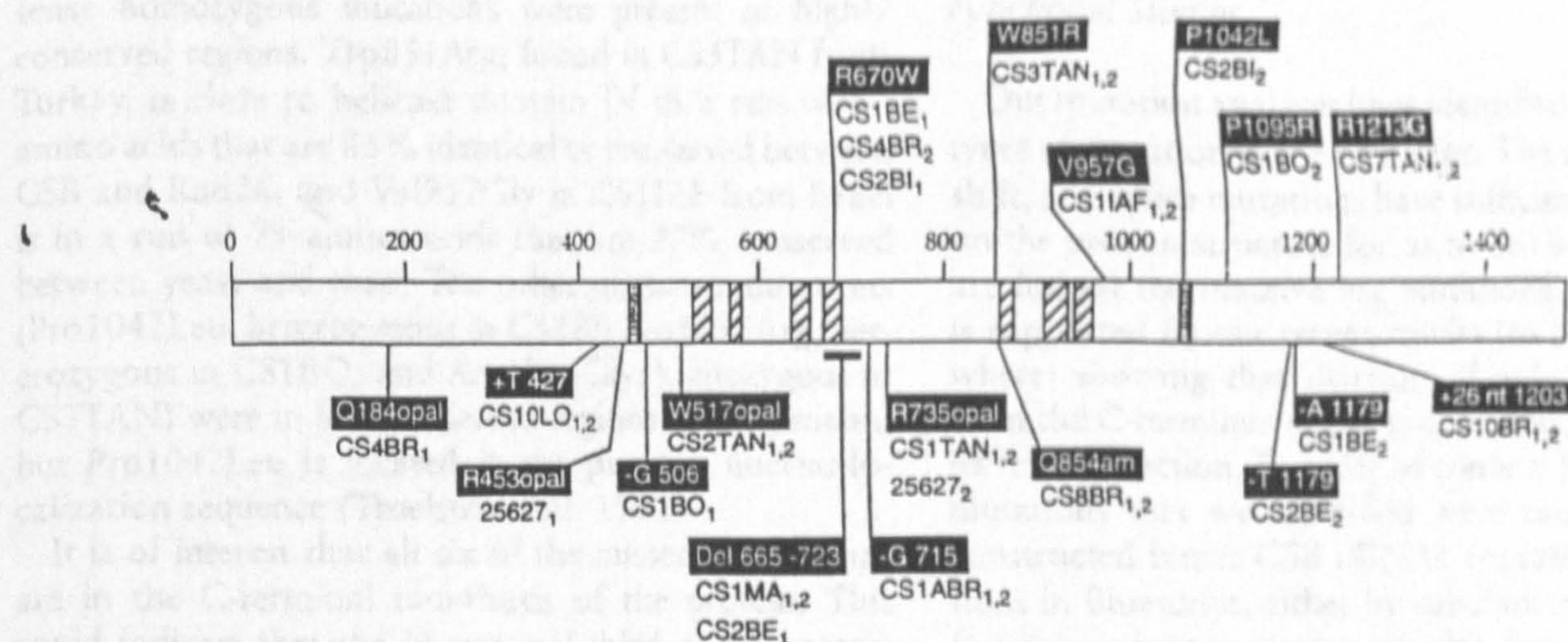


Figure 1 Inactivating amino acid changes caused by mutations in the CSB gene. The CSB protein is shown with the seven domains conserved in DNA helicases (*hatched boxes*) and with the putative nuclear-localization signals (*gray-shaded boxes*). Amino acid changes resulting from mutations are shown boxed, with the change indicated by white letters on a black background, and the cell-line designations are shown below them. Subscripts "1" and "2" denote the different alleles. Mutations shown below the depicted protein result in protein truncations, whereas those above the depicted protein are all single-amino-acid changes.

mined the sites of the mutations in cells from his similarly affected brother, coded as "25627." The patient is a compound heterozygote: one allele contains the Arg735opal mutation found in CS1TAN, and the other allele contains the mutation C1436T, resulting in Arg453opal. CS4BR, derived from a patient in the United Kingdom, is a compound heterozygote with the mutation C629T, resulting in Gln184opal in one allele. CS1BO is a rare example of a black patient with CS. She is a compound heterozygote with C2918T in one allele, resulting in Arg947opal; this mutation is, however, not relevant for the pathological phenotype, since it is 1,321 bases downstream of a frameshift mutation in the same allele (see below).

Frameshifts and Insertions

Single-base frameshifts, all resulting in stop codons 7–35 amino acids downstream, were detected in six patients and are shown in table 3. Three of these were single-base deletions or insertions in runs of five or six

identical bases, presumably the result of replication slippage, which could also account for the loss of an A residue sandwiched between a run of 5 T's and a single T in CS1BE. In CS1BO there was a loss of a G residue from the center of a 12 bp inverted repeat. In CS10BR there was an insertion of 26 bases at position 3686, in the middle of exon 18. The origin of this insertion is not known. At the time of the writing of this report, this sequence was not present in any of the publicly available sequence databases.

Missense Mutations

A missense mutation, C2087T, causing an Arg→Trp change at amino acid 670, has been found in one allele of three unrelated patients, CS4BR and CS2BI from the United Kingdom and CS1BE (GM1629) from the United States. Arg670 is in helicase domain III and is part of a stretch of 34 amino acids that are 97% identical or conserved in CSB and the *Saccharomyces cerevisiae* homologue, Rad26 (van Gool et al. 1994). Two other mis-

Table 3

Frameshifts and Insertions

Cell Strain	Status	Mutation	First Mutated Codon	Sequence	Size of Truncated Protein (amino acids)
CS10LO	Homozygous	+T 1355–1359	427	ACTTTTTTCC	434
CS1ABR	Homozygous	–G 2218–2223	715	ATGGGGGGGAT	737
CS2BE	Heterozygous	–T 3610–3614	1179	AATTTTTTAT	1,199
CS1BE	Heterozygous	–A 3615	1179	TTTTTATAAGC	1,199
CS1BO	Heterozygous	–G 1597	506	AAAAAGCTTTTTT	541
CS10BR	Homozygous	+26nt ^a 3686	1203	ACCAA +26 nt ^a +AGCAA	1,234

^a Insertion of GG GCT GGC TGC TTA AGG TCC ACC TTA.

sense homozygous mutations were present in highly conserved regions. Trp851Arg, found in CS3TAN from Turkey, is close to helicase domain IV in a run of 37 amino acids that are 83% identical or conserved between CSB and Rad26, and Val957Gly in CS1LAF from Israel is in a run of 79 amino acids that are 97% conserved between yeast and man. The other missense mutations (Pro1042Leu, heterozygous in CS2BI; Pro1095Arg, heterozygous in CS1BO; and Arg1213Gly, homozygous in CS7TAN) were in less-conserved regions of the protein, but Pro1042Leu is located in the putative nuclear-localization sequence (Troelstra et al. 1992).

It is of interest that all six of the missense mutations are in the C-terminal two-thirds of the protein. This could indicate that the N-terminal third of the protein is less important for the repair function of the CSB protein and that this domain is involved in some other function (for other possible roles of the CSB protein, see the Discussion section). Preliminary results from site-directed mutagenesis of the N-terminal part of the gene are consistent with this possibility.

Splice Mutations

In two patients the CSB mRNA was abnormally spliced. Exon 10 was absent from the cDNA of CS1MA, suggesting that the mutation in both alleles affected the splicing of this exon. This same exon was missing in one allele of CS2BE. We have sequenced the regions around the splice-donor site of intron 10 and the splice-acceptor site of intron 9, in genomic DNA of CS1MA, but this did not reveal the cause of the abnormal splicing. The loss of exon 10 results in an in-frame deletion of 59 amino acids encompassing the whole of helicase domain III.

Polymorphisms

In addition to the inactivating mutations, we have detected the following changes in at least two patients: G214C (Leu45), C2830T (Gly917), G1275A (Gly399Asp), A3368G (Met1097Val), and A4317G (Gln1413Arg). We presume that Gly399Asp and Met1097Val are silent. Residue 399 is Gly in the human gene but is Asp in the mouse gene, and Met1097Val is a conservative change. We show below that Gln1413Arg is not inactivating. In patient CS7TAN, we found three of the above-mentioned polymorphic changes, as well as Arg1213Gly, which we show to be inactivating (see below), and Lys255Thr, all mutations being homozygous. Since Arg1213Gly is sufficient to account for the repair-deficient phenotype, we presume that Lys255Thr does not contribute to the phenotype.

Functional Studies

Our mutation analyses have identified several different types of mutation in the CSB gene. The nonsense, frameshift, and splice mutations have sufficiently severe effects on the protein structure for us to be confident that they are indeed the inactivating mutations. This contention is supported by our recent results (to be presented elsewhere) showing that deletion of only 25 amino acids from the C-terminus of the protein completely abolishes its repair function. In order to confirm that the missense mutations that we identified were causative, we have constructed intact CSB cDNAs containing these mutations in Bluescript, either by subcloning fragments used in our sequence analysis or by site-directed mutagenesis. All cloned PCR products or mutated plasmids were sequenced in their entirety, to ensure that the DNA contained only the desired mutations. The CSB cDNA was then transferred, as a *Bam*HI fragment, from Bluescript into one of the mammalian expression vectors pCDNA3 or pCIneo. Normal and mutant plasmids were transfected into the UV-sensitive CHO cell line UV61 (from the same complementation group as the CS-B cells), and selection for the *neoR* gene was applied by use of G418. Clones were picked and expanded, and the remaining clones in the dishes were trypsinized, respread, and subjected to UV-irradiation selection with three doses, each of 4 Jm⁻². This gave a crude measure of UV sensitivity of the transfectants. Whereas a confluent layer of cells resulted from the culture of UV61 cells that were transfected with the wild-type CSB gene and then UV irradiated, very few cells survived after transfection with either empty vector or the CSB gene containing the mutations R670W, W851R, V957G, P1042L, or R1213G. In contrast, when the cDNA contained mutation Q1413R, which we presumed to be a silent polymorphism, the cells survived the UV-irradiation selection, as did the wild-type transfectants.

The clones that were picked and expanded without UV irradiation were examined by PCR, to ensure that the transfected human CSB cDNA was intact. For each transfection, two clones in which PCR indicated that the whole of the cDNA was present were examined quantitatively for the effects that UV irradiation had on cell survival. Results are shown in figure 2. These results confirmed the findings seen with the batch cultures described above. In early experiments, in which the pCDNA3 expression vector was used, clones from cells transfected with the wild-type gene were much more resistant than were untransfected UV61 cells, although the survival did not reach that of the wild-type parental cell line AA8. In later experiments, with the pCIneo vector, the wild-type CSB gene restored the survival of UV61 cells up to the level of the parental AA8 line. Irrespective of which vector was used, the survival of

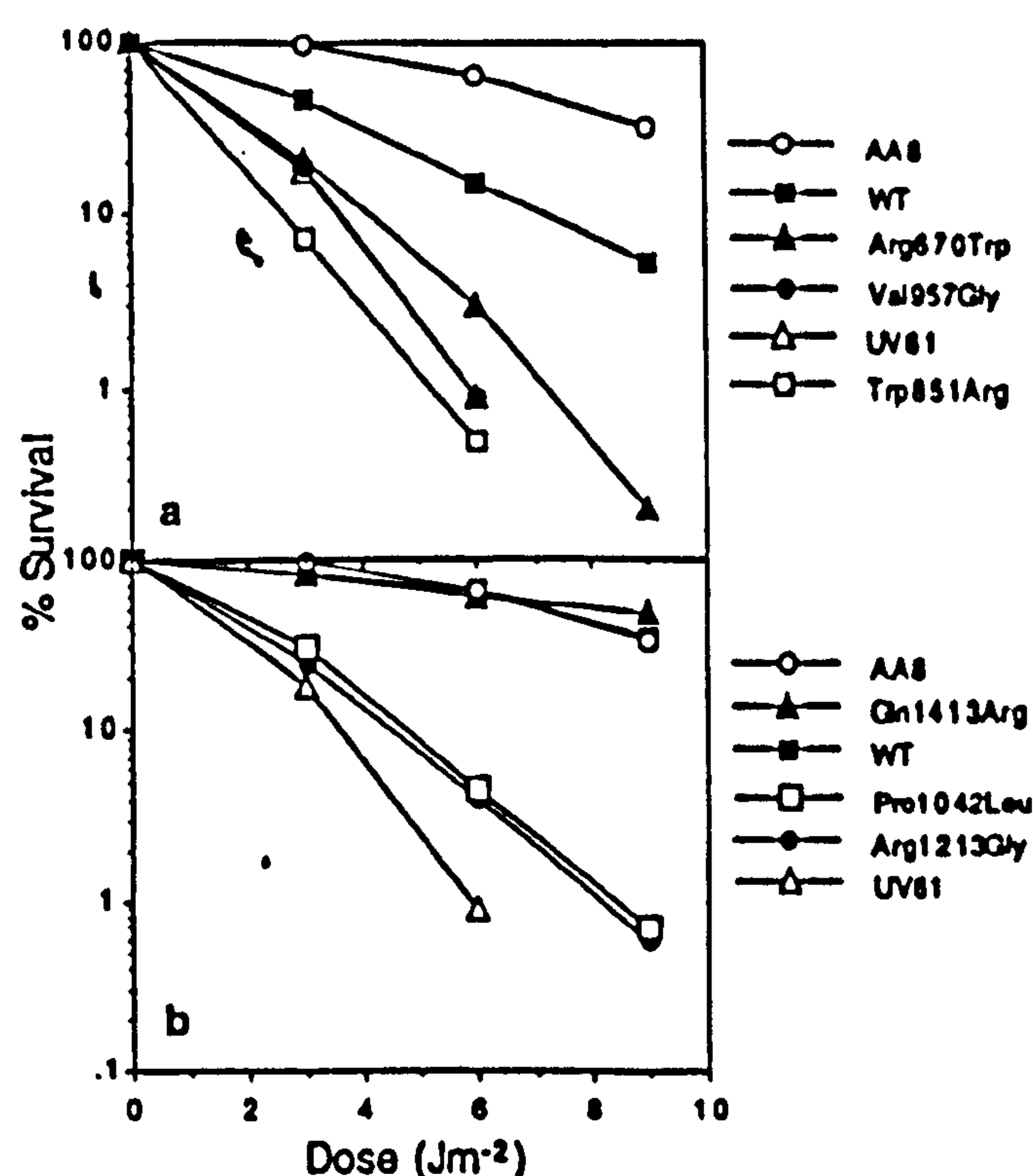


Figure 2 UV-irradiation survival of UV61 cells transfected with the CSB gene containing different mutations. Transfectants are from CSB constructs made in pCDNA3 (a) or in pCIneo (b). "AA8" denotes the normal parental hamster cell line from which UV61 was derived, and all other cell lines are UV61 cells transfected with CSB constructs containing the indicated mutations.

UV61 cells transfected with DNA containing the mutations that we had identified as inactivating was indistinguishable from that of cells transfected with the empty vector. In contrast, the survival of transfectants with the Q1413R DNA approached that of the normal transfectant.

Discussion

We have identified the inactivating mutations in the CSB gene in 16 affected patients. The results are summarized in figure 1 and tables 2 and 4. A variety of different types and positions of mutations result in the CS phenotype. Of the 18 identified inactivating mutations, four were CG→TA transitions at CpG sites, resulting from deamination of 5-methylcytosine, a proportion similar to that found in other human genetic disorders; and a further five were transitions at other sites, two were transversions, six were frameshifts, and one was an aberrant splicing.

Twelve of the 18 mutations resulted in severely truncated products, because of either stop codons, frameshifts, or splice abnormalities. In nine patients (table 4, top), both alleles were affected in this way, and it is

highly unlikely that the CSB protein in these individuals would have any functional ability. This was also found in the single patient analyzed by Troelstra et al. (1992). This demonstrates that CSB is not an essential gene, as also has been found for the yeast RAD26 homologue (van Gool et al. 1994). Further evidence to support this comes from the generation of a CSB knockout mouse, which showed near-normal development (van der Horst et al. 1997). Of the six missense mutations, three were located in regions that were highly conserved not only in the mouse CSB protein but also in the homologous yeast Rad26 protein. These mutations lie either within or very close to the helicase domains of the protein and are likely to abolish any function associated with these domains. Of the three other mutated sites, Arg1213 is conserved in the human and mouse proteins. Pro1042, which is mutated to leucine in one allele of CS2BI, is not present in the mouse protein (G. T. J. van der Horst and J. H. J. Hoeijmakers, personal communication), but it is located in a putative nuclear-localization signal, and we have shown that this alteration, like the other amino acid changes indicated in figure 1, fails to restore UV-irradiation resistance to UV61 cells.

The CS patients studied come from a wide range of racial backgrounds, including Caucasian, Turkish, Indo-Pakistani, Israeli Jewish, and Black patients. The clinical features of CS are fairly heterogeneous, as can be seen in table 4, in our earlier clinical survey (Lehmann et al. 1993), and in reviews in the literature (Nance and Berry 1992). We previously had found that, in those patients in whom RNA synthesis failed to recover after UV irradiation (a feature that we use as a diagnostic marker for CS), the magnitude of this defect was broadly similar, irrespective of the clinical features (Lehmann et al. 1993). Similar observations have been made in Japanese patients, by Sugita et al. (1991). Lowry (1982 [also see Nance and Berry 1992]) has suggested that CS can be divided clinically into type I, with classical CS symptoms that become manifest within the first few years of life, and type II, with more-severe symptoms already manifest prenatally. Cellular defects in the response to UV irradiation appear to be similar for type I patients and type II patients (Sugita et al. 1991). We do not have full clinical histories of all the patients in our study, but at least four of them—CS1MA, CS1ABR, CS8BR, and CS1BO—had the more-severe, type II features. Three of these are homozygous for mutations resulting in severely truncated proteins, whereas CS1BO is a compound heterozygote and has the Pro1095Arg change in one allele. However, six other patients also produced truncated products from both alleles but had classical, type I features. It thus seems unlikely that the assignment into one of the two clinical types can be based on the position or type of mutation; the clinical disparity is more likely

Table 4

CS Mutations and Symptoms

Patient ^a	Amino Acid Change ^b	Racial Group	Age at Biopsy (years)	Age at Onset (years)	Pigmentary Retinopathy Symptoms	Gait Defect ^c	Cataracts	Caries	Reference ^d
Truncated protein in both alleles:									
CS10LO	+1 nt 427	Black	4	.5	No	Yes	No	?	
CS2TAN	Trp517stop	Turkish	9	1.5	Yes	Yes	No	No	Lehmann et al. (1993); 88
CS1MA*	Deletion of 665-723	Caucasian	4 (died)	0	?	Yes	Yes	Yes	Lehmann et al. (1993); 34
CS1ABR*	-1 nt 715	Caucasian	6	0			Yes		Stefanini et al. (1996)
CS1TAN	Arg735stop	Turkish	8	1.5	Yes	Yes	No	Yes	Lehmann et al. (1993); 77
CS2627	Arg735stop	Caucasian	5		Yes	Yes			
	Arg453stop								
CS8BR*	Gln854stop	Indo-Pakistani	.5	0					
CS2BE	-1 nt 1178	Caucasian	21	1	Yes	Yes			Brumback et al. (1978)
	Deletion of 665-723								
CS10BR	+26 nt 1203	Caucasian	5	.2	No	Yes	Yes	No	
Missense mutation in at least one allele:									
CS4BR	Arg670Trp	Caucasian	24	3	Yes	Yes	Yes	Yes	Lehmann et al. (1993); 39
	Gln184stop								
CS2BI	Arg670Trp	Caucasian	16	.7	Yes	Yes			Lehmann et al. (1993); 9
	Pro1042Leu								
CS1BE	Arg670Trp	Caucasian	11	1.5	Yes	Yes	No		Brumback et al. (1978)
	-1 nt 1179								
CS3TAN	Trp851Arg	Turkish	1.7	.3	No	Yes	No	No	Lehmann et al. (1993); 98
CS1IAF	Val957Gly	Israeli Jewish	7	2	Yes			Yes	Lehmann et al. (1993); 67
CS1BO*	Pro1095Arg	Black	.75	0	Yes	NA	Yes	Yes	Stefanini et al. (1996)
	-1 nt 506								
CS7TAN	Arg1213Gly	Turkish	18	2	Yes	Yes	No	Yes	

* Listed in order of amino acid position of truncation or frameshift.

^b The mutation responsible for the pathological phenotype is underlined; for compound heterozygotes, the less severe mutation is underlined. This is assumed to be the more C-terminal of two truncating mutations—or a missense mutation, if one allele is a truncation and the other is a missense mutation.

^c NA = not applicable.

^d Numbers refer to the designations of the patients in Lehmann et al. (1993).

* Type II patient with symptoms at birth.

to result from other differences in genetic background or in the intrauterine environment.

CS-B (and CS-A) cells are sensitive to the lethal effects of UV irradiation and of a variety of chemical carcinogens (Wade and Chu 1979). This is associated, in all cases, with a failure of RNA synthesis to recover after DNA damage (Mayne and Lehmann 1982). Until recently this had been attributed to a deficiency in the ability of CS to perform TCR—that is, the rapid removal of damage from the transcribed strand of active genes (van Hoffen et al. 1993). These observations, in turn, have led to the suggestion that the function of the CS gene products is to recruit the NER machinery to actively transcribed regions of DNA, a role analogous to that of the transcription-repair coupling factor Mfd protein in *Escherichia coli* (Selby and Sancar 1993). Although this model has been widely accepted, it has recently been questioned by some of its original proponents, who found that, after damage to CS cells by the carcinogen N-acetoxy-acetylaminofluorene, TCR occurs at normal levels, although RNA synthesis fails to recover (van Oosterwijk et al. 1996). Those authors proposed that, after UV irradiation, the dual-function transcription/repair factor TFIIH is converted into a form utilized in NER and that the function of the CS proteins is subsequently to revert TFIIH from repair mode to transcription mode, so that transcription can recommence. It has been proposed that, in CS patients, TFIIH is locked in the repair mode and transcription fails to recover (van Oosterwijk et al. 1996). In an alternative model, the features of CS have been suggested to result from defects in base excision repair of oxidative base damage (Cooper et al. 1997).

Except for sun sensitivity, the clinical features of CS cannot be obviously attributed to defects in DNA repair. This is in contrast to xeroderma pigmentosum, in which most of the symptoms are readily explicable in terms of a repair deficiency. It has therefore been suggested that the CSB protein has a second function, possibly an involvement with transcription, as has been found to be the case with TFIIH (Bootsma and Hoeijmakers 1993; Friedberg 1996). Although this is an attractive hypothesis, CSB is not an essential gene, and therefore the CSB protein cannot be an integral part of the transcriptional machinery. Evidence in support of a role for the CS proteins in general transcription comes both from data suggesting that transcription rates are reduced in CS cells (Balajee et al. 1997) and from data showing that 10%–15% of RNA polymerase II is strongly associated with CSB protein in cell extracts (van Gool et al. 1997a). This has led to the suggestion that the role of CSB might be to release stalled transcription, by direct action on the RNA polymerase II (van Gool et al. 1997b). An alternative possibility is that CS genes are involved in tissue-specific transcription. A gene that has been des-

ignated "ATR-X" is another member of the SWI/SNF superfamily. It is an X-linked gene that, when mutated, specifically down-regulates expression of several genes, including the α -globin gene on chromosome 16 (Gibbons et al. 1995; Picketts et al. 1996). One can envisage that CSB might act in a similar manner on the expression of a set of tissue-specific genes involved in aspects of development that are abnormal in CS patients.

Acknowledgments

The work was supported in part by a MRC GAHH supplementation grant to A.R.L. and D.L.M., by CEC Human Capital and Mobility Grant CHRXCT940443 to A.R.L. and M.S., and by Telethon-Italy grant E.550 to M.S.

References

- Arlett CF, Lehmann AR, Giannelli F, Ramsay CA (1978) A human subject with a new defect in repair of ultraviolet damage. *J Invest Dermatol* 70:173–177
- Balajee AS, May A, Dianov GL, Friedberg EC, Bohr VA (1997) Reduced RNA polymerase II transcription in intact and permeabilized Cockayne syndrome group B cells. *Proc Natl Acad Sci USA* 94:4306–4311
- Bootsma D, Hoeijmakers JHJ (1993) Engagement with transcription. *Nature* 363:114–115
- Brumback RA, Yoder FW, Andrews AD, Peck GL, Robbins JH (1978) Normal pressure hydrocephalus: recognition and relationship to neurological abnormalities in Cockayne's syndrome. *Arch Neurol* 35:337–345
- Cairns BR, Lorch Y, Li Y, Zhang M, Lacomis L, Erdjument-Bromage H, Tempst P, et al (1996) RSC, an essential, abundant chromatin-remodeling complex. *Cell* 87:1249–1260
- Cooper PK, Nousepikel T, Clarkson SG, Leadon SA (1997) Defective transcription-coupled repair of oxidative base damage in Cockayne syndrome patients from XP group G. *Science* 275:990–993
- Friedberg EC (1996) Cockayne syndrome—a primary defect in DNA repair, transcription, both or neither? *BioEssays* 18: 731–738
- Gibbons RJ, Picketts DJ, Villard L, Higgs DR (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with α -thalassemia (ATR-X syndrome). *Cell* 80:837–845
- Henning KA, Li L, Iyer N, McDaniel LD, Reagan MS, Legerski R, Schultz RA, et al (1995) The Cockayne-syndrome group-A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA-polymerase-II TFIIH. *Cell* 82:555–564
- Hoeijmakers JHJ (1993) Nucleotide excision repair II: from yeast to mammals. *Trends Genet* 9:211–217
- Hoeijmakers JHJ, Egly J-M, Vermeulen W (1996) TFIIH: a key component in multiple DNA transactions. *Curr Opin Genet Dev* 6:26–33
- Kunkel TA, Roberts JD, Zakour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol* 154:367–382

- Lehmann AR (1982) Three complementation groups in Cockayne syndrome. *Mutat Res* 106:347-356
- Lehmann AR (1995) Nucleotide excision repair and the link with transcription. *Trends Biochem Sci* 20:402-405
- Lehmann AR, Bootsma D, Clarkson SG, Cleaver JE, McAlpine PJ, Tanaka K, Thompson LH, et al (1994) Nomenclature of human DNA repair genes. *Mutat Res* 315:41-42
- Lehmann AR, Thompson AF, Harcourt SA, Stefanini M, Norris PG (1993) Cockayne's syndrome: correlation of clinical features with cellular sensitivity of RNA synthesis to UV-irradiation. *J Med Genet* 30:679-682
- Lowry RB (1982) Early onset of Cockayne syndrome. *Am J Med Genet* 13:209-210
- Mayne LV, Lehmann AR (1982) Failure of RNA synthesis to recover after UV-irradiation: an early defect in cells from individuals with Cockayne's syndrome and xeroderma pigmentosum. *Cancer Res* 42:1473-1478
- Nance MA, Berry SA (1992) Cockayne syndrome: review of 140 cases. *Am J Med Genet* 42:68-84
- Pazin MJ, Kadonaga JT (1997) SWI2/SNF2 and related proteins: ATP-driven motors that disrupt protein-DNA interactions. *Cell* 88:737-740
- Peterson CL, Tamkun JW (1995) The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem Sci* 20:143-146
- Picketts DJ, Higgs DR, Bachoo S, Blake DJ, Quarell OWJ, Gibbons RJ (1996) ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Hum Mol Genet* 5:1899-1907
- Selby CP, Sancar A (1993) Molecular mechanism of transcription repair coupling. *Science* 260:53-58
- (1997) Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *J Biol Chem* 272:1885-1890
- Stefanini M, Fawcett H, Botta E, Nardo T, Lehmann AR (1996) Genetic analysis of twenty-two patients with Cockayne syndrome. *Hum Genet* 97:418-423
- Steingrimsdottir H, Rowley G, Waugh A, Beare D, Ceccherini I, Cole J, Lehmann AR (1993) Molecular analysis of mutations in the hprt gene in circulating lymphocytes from normal and DNA-repair-deficient donors. *Mutat Res* 294:29-41
- Sugita K, Takanashi J, Suzuki N, Niimi H (1991) Comparison of cellular sensitivity to UV killing with neuropsychological impairment in Cockayne syndrome patients. *Brain Dev* 13:163-166
- Tanaka K, Kawai K, Kumahara Y, Ikenaga M (1981) Genetic complementation groups in Cockayne syndrome. *Somat Cell Genet* 7:445-456
- Troelstra C, Odijk H, de Wit J, Westerveld A, Thompson LH, Bootsma D, Hoeijmakers JHJ (1990) Molecular cloning of the human DNA repair gene ERCC6. *Mol Cell Biol* 10:5806-5813
- Troelstra C, van Gool A, de Wit J, Vermeulen W, Bootsma D, Hoeijmakers JHJ (1992) ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. *Cell* 71:939-953
- van der Horst GTJ, van Steeg H, Berg RJW, van Gool AJ, de Wit J, Weeda G, Morreau H, et al (1997) Defective transcription-coupled repair in Cockayne syndrome B mice is associated with skin cancer predisposition. *Cell* 89:425-435
- van Gool AJ, Citterio E, Rademakers S, van Os R, Vermeulen W, Constantinou A, Egly J-M, et al (1997a) The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in a RNA polymerase II containing complex. *EMBO J* 16:5955-5965
- van Gool AJ, van der Horst GTJ, Citterio E, Hoeijmakers JHJ (1997b) Cockayne syndrome: defective repair of transcription. *EMBO J* 16:4155-4162
- van Gool AJ, Verhage R, Swagemakers SMA, van de Putte P, Brouwer J, Troelstra C, Bootsma D, et al (1994) RAD26, the functional *S. cerevisiae* homolog of the Cockayne syndrome B gene ERCC6. *EMBO J* 13:5361-5369
- van Hoffen A, Natarajan AT, Mayne LV, van Zeeland AA, Mullenders LHF, Venema J (1993) Deficient repair of the transcribed strand of active genes in Cockayne's syndrome cells. *Nucleic Acids Res* 21:5890-5895
- van Oosterwijk MF, Versteeg A, Filon R, van Zeeland AA, Mullenders LHF (1996) The sensitivity of Cockayne's syndrome cells to DNA-damaging agents is not due to defective transcription-coupled repair of active genes. *Mol Cell Biol* 16:4436-4444
- Wade MH, Chu EHY (1979) Effects of DNA damaging agents on cultured fibroblasts derived from patients with Cockayne syndrome. *Mutat Res* 59:49-60